Earth Surf. Dynam., 13, 1–22, 2025
https://doi.org/10.5194/esurf-13-1-2025

Earth Surface Dynamics
Open Access
EGU

Examination of analytical shear stress predictions for coastal dune evolution

Orie Cecil[1], Nicholas Cohn[2], Matthew Farthing[1], Sourav Dutta[1,3], and Andrew Trautz[4]

[1]Coastal and Hydraulics Laboratory, US Army Engineer Research and Development Center, 3909 Halls Ferry Road, Vicksburg, MS 39180, USA
[2]Coastal and Hydraulics Laboratory, US Army Engineer Research and Development Center, 1261 Duck Road, Duck, NC 27949, USA
[3]Oden Institute for Computational Engineering and Sciences, University of Texas at Austin, 201 E. 24th Street, Austin, TX 78712, USA
[4]Geotechnical and Structures Laboratory, US Army Engineer Research and Development Center, 3909 Halls Ferry Road, Vicksburg, MS 39180, USA

Correspondence: Orie Cecil (orie.m.cecil@erdc.dren.mil)

Received: 21 March 2024 – Discussion started: 4 April 2024
Revised: 2 August 2024 – Accepted: 21 October 2024 – Published: 2 January 2025

Abstract. Existing process-based models for simulating coastal foredune evolution largely use the same analytical approach for estimating wind-induced surface shear stress distributions over spatially variable topography. Originally developed for smooth, low-sloping hills, these analytical models face significant limitations when the topography of interest exhibits large height-to-length ratios and/or steep, localized features. In this work, we utilize computational fluid dynamics (CFD) to examine the error trends of a commonly used analytical shear stress model for a series of idealized two-dimensional dune profiles. It is observed that the prediction error of the analytical model increases compared to the CFD simulations for increasing height-to-length ratio and localized slope values. Furthermore, we explore two data-driven methodologies for generating alternative shear stress prediction models, namely, symbolic regression and linear, projection-based, non-intrusive reduced-order modeling. These alternative modeling strategies demonstrate reduced overall error but still suffer in their generalizability to broader sets of dune profiles outside of the training data. Finally, the impact of these improvements on aeolian sediment transport fluxes is examined to demonstrate that even modest improvements to the shear stress prediction can have significant impacts on dune evolution simulations over engineering-relevant timescales.

1 Introduction

Complex landforms are common features in sandy, subaerial environments arising out of spatial sediment transport gradients. Parabolic dunes, for example, develop at sites with strong, unidirectional winds where spatially nonuniform vegetation stabilization can trigger a feedback cycle generating a U-shaped morphological feature (Yan and Baas, 2015). Similarly, barchan dunes are three-dimensional landforms arising out of nonuniform sediment supply that initiates the feature and is reinforced through flow–sediment interactions. Numerous other types of landforms exist depending on wind speed and directionality, hydrologic properties, local sediment supply and bed characteristics, vegetation effects, and the presence of upwind obstacles. The ability to predict the formation and evolution of such landforms is critical in many sandy subaerial regions due to infrastructural and transportation risks associated with migrating sand (e.g., Bruno et al., 2018; Khalaf and Al-Ajmi, 1993; Puy et al., 2018). This applied management need has partially driven the development of a wide range of tools of various spatial resolutions, process capabilities, and computational needs for synthesizing components of the sediment transport processes and morphodynamic feedbacks contributing to aeolian landform develop-

ment (Diniega et al., 2010; Durán et al., 2010; Keijsers et al., 2016; Luna et al., 2011; Sauermann et al., 2001).

The need for quantitative, predictive wind-driven morphological tools in the coastal zone, where small modifications to the topography have important implications for flooding-related hazards during storms (e.g., Figlus, 2022; Hanley et al., 2014), is particularly pressing (Elko et al., 2016). Specifically, coastal foredunes often represent the first line of defense for such hazards and are therefore increasingly being encouraged to grow or are being constructed in order to add resilience to the system. However, beach-dune systems are also commonly characterized by abrupt spatial transport gradients due to moisture effects on the transport field, sediment heterogeneity, vegetation, and steep topography. Over the past decade, foundational tools for simulating mesoscale landform development in desert environments have been ported to coastal environments, in part through the incorporation of vegetation–shear stress interactions (Durán and Moore, 2013), time evolution of the vegetation field (Charbonneau et al., 2022), improved representation of multi-fraction transport effects (Hoonhout and de Vries, 2016), and enhanced representations of groundwater and surface moisture effects (Hage et al., 2020; Hallin et al., 2023).

While new process capabilities are critical for improved representations of transport processes in supply-limited systems (van IJzendoorn et al., 2023) and ecological effects controlling deposition patterns (Dickey et al., 2023; Okin, 2008), many of these process-based tools being developed for coastal dune evolution utilize numerical representations of the same analytical solutions to account for topographic effects on wind-related spatial shear stress perturbations derived from Kroy et al. (2002) and related works. For example, the Coastal Dune Model includes a 2D bidirectional wind solving capability based on the Kroy et al. (2002) solution (hereby referred to as KSH) that has shown the ability to account for flow–sediment–morphology interactions related to the building of vegetated coastal foredunes and foredune ridges (Duran and Moore, 2013; Moore et al., 2016). Similarly, a 1D representation of this approach has been incorporated into the Duna Model (Roelvink and Costas, 2019), and a wind-aligned grid rotation scheme has allowed 2D omnidirectional winds for characterizing wind flow perturbations in AeoLiS (van Westen, 2018).

Although the implementation of these routines differs, all widely used process-based coastal dune growth models operate from the same foundational assumptions and numerical approximations for shear stress perturbations, as described in more detail in Sect. 1.2. However, at sites characterized by often steep topographic gradients, the validity of these existing numerical expressions for topographically induced flow (shear stress) acceleration patterns and separation bubble behavior has not been widely assessed. As such numerical tools move from mostly theoretical applications (Parteli et al., 2014; Durán Vinent and Moore, 2015) to applications of real-world dune dynamics (Kombiadou et al., 2023; Strypsteen and de Vries, 2023; van Westen et al., 2024), constraining errors in representations of physical processes becomes particularly important. Dune systems in particular have a broad range of forms and sizes depending on the vegetation type and species (Zarnetske et al., 2012), sediment supply (Psuty, 2008), and disturbance history (Robin et al., 2021), as partially represented by the wide range in dune height (H) to toe-crest length (L_{base}) ratio shown in Fig. 1. Given that mesoscale morphological changes are often the result of small sediment transport gradients that aggregate over extended timescales, an incomplete representation of these complex wind dynamics limits the ability to successfully predict dune evolution across relevant engineering timescales (hours to decades) – posing limitations for quantifying future risk from encroachment and increased flooding potential, as well as engineering solutions to limit these hazards.

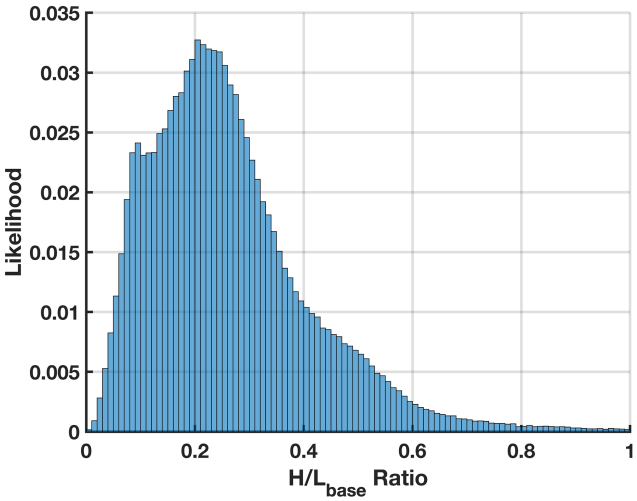

Figure 1. Distribution of H/L_{base} for coastal foredunes along the outer coast of the continental United States derived from Mull and Ruggiero (2014) and Doran et al. (2017), as compiled in Cohn (2022), where L_{base} is assumed to be the length scale from the dune toe to the dune crest.

In this work we aim to examine the limits and error trends of analytical bed shear stress predictions by comparing to computational fluid dynamics (CFD) simulations for a range of idealized dune profiles. Furthermore, we explore alternative data-driven models through symbolic regression (SR) methods and projection-based, non-intrusive, reduced-order modeling (NIROM) techniques. Our goal is to provide insight into the errors incurred by and subsequent sediment transport effects of using traditional theoretical bed shear stress predictions in dune evolution models, as well as to explore the advantages and disadvantages of some possible alternative flow modeling strategies. We begin by providing a brief overview of analytical bed shear stress predictions, the use of CFD in modeling flow dynamics over

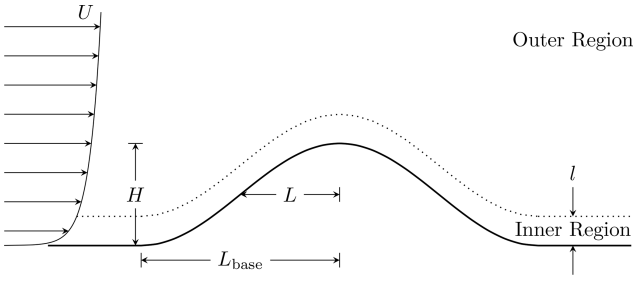

Figure 2. Sketch depicting the important geometric dimensions and regions for flow over simplified dune profiles.

dunes, symbolic regression methods, and projection-based NIROM techniques in Sect. 2. In Sect. 3, we lay out our approach for using these methods to produce alternative bed shear stress predictions. Results of CFD, symbolic regression, and NIROM predictions and their comparison to KSH are examined in Sect. 4. Finally, an intercomparison of the methods considered and the subsequent impact on dune evolution models are discussed in Sect. 5.

2 Background

2.1 Analytical solutions for flow over low sloping dunes

One of the earliest analytical models for predicting the near-surface flow over a low-sloping hill was developed by Jackson and Hunt (1975). Given the overall height, H, and the characteristic length, L, defined as the half-length at half-height, the slope is assumed to be on the order of the height-to-length ratio, H/L, which is assumed to be small. Additionally, the characteristic length-to-roughness-length ratio, L/z_0, which is assumed to be large and uniform, also plays a central role in the analysis. Jackson and Hunt (1975) split the atmospheric boundary layer into two regions where the outer layer is treated as inviscid, while flow in the inner layer is driven by the induced pressure gradient and turbulence effects. The various important dimensions and regions are depicted in Fig. 2. The Reynolds-averaged Navier–Stokes (RANS) equations are then linearized and a solution obtained through a series of scaling arguments and matched asymptotic expansions.

Since the initial work of Jackson and Hunt (1975), the theory has been extended to three dimensions (Mason and Sykes, 1979), upstream velocity profiles with strong shear (Hunt et al., 1988), and higher-order corrections (Weng et al., 1991). However, the normalized bed shear stress perturbation remains essentially the same and is given by Weng et al. (1991) as

$$\tau' = \frac{\tau}{\tau_0} - 1 \tag{1}$$

$$\mathcal{F}[\tau'] = \frac{2H/L}{U^2(l)}\mathcal{F}[\sigma']\left[1 + \frac{2\ln|kL| + 4\gamma + 1 + i\pi}{\ln(l/z_0)}\right] \tag{2}$$

$$\sigma' = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{f'(\xi')}{\xi - \xi'}\mathrm{d}\xi', \tag{3}$$

where σ' is the leading order normalized pressure perturbation given by the Hilbert transform of the hill slope, z_0 the surface roughness length, k the Fourier wave number, $\xi = x/L$ a normalized streamwise coordinate, $\gamma \approx 0.57721$ Euler's constant, and l the length scale of the inner region given by

$$l\ln\left(\frac{l}{z_0}\right) = 2\kappa^2 L. \tag{4}$$

More recently, Kroy et al. (2002) further simplified the expression for a transverse dune, leading to

$$\mathcal{F}[\tau'] = A(|k| + iBk)\mathcal{F}[h(x)] \tag{5a}$$

$$A = \frac{\ln(\Phi^2/\ln\Phi)^2}{2(\ln\phi)^3}[1 + \ln\phi + 2\ln(\pi/2) + 4\gamma] \tag{5b}$$

$$B = \pi[1 + \ln\phi + 2\ln(\pi/2) + 4\gamma]^{-1} \tag{5c}$$

$$\phi \equiv 2\kappa^2\Phi/\ln\phi \tag{5d}$$

$$\Phi = L/z_0, \tag{5e}$$

where $h(x)$ represents the dune topography and $\kappa = 0.41$ is the von Kármán constant. It should be noted that we have retained the terms involving the logarithmic dependence on the characteristic length which stems from Eq. (2), i.e., $2\ln(\pi/2)$, in contrast to the description in Kroy et al. (2002), which indicates that these terms are neglected in their subsequent analysis.

As alluded to earlier, three inherent limitations exist with these asymptotic-based analytical models when applying them to realistic coastal dunes. First, linearization of the underlying equations precludes the ability to capture separated flow effects. Traditionally, this has been overcome by extending the dune surface with an empirical separation bubble profile (Kroy et al., 2002; Schatz and Herrmann, 2006) or by applying a nonlinear correction to the predicted bed shear stress based on numerical simulations of forward- and backward-facing steps (Pelletier, 2009). However, it should be noted that these ad hoc extensions do not account for the dependence of the separating streamline and reattachment angle on the upwind shear velocity (Araújo et al., 2013). Additionally, by setting the shear stress to zero in these regions the potential for reverse transport in the lee is overlooked. Secondly, the assumption of a small H/L and dune slope can lead to issues when applying the model to dunes with

sharp local features, such as scarps, that may otherwise satisfy the H/L criteria and not exhibit separation on the lee side. Again, while the shear stress perturbation can be limited to $\tau' \geq -1$ in the case of separation at a scarp (Durán Vinent and Moore, 2015), the presence of this locally steep topography influences the shear stress across the crest and in the lee of the dune as well (Bauer and Wakes, 2022), leading to additional errors. Additionally, the majority of dunes, at least in the coastal domain (e.g., Fig. 1), do not conform to the low-slope assumption, limiting the use of these analytical approaches to a small subset of morphological dune types. Finally, while the surface roughness, z_0, is assumed to be constant over the dune, realistic coastal dunes rarely exhibit a uniform surface roughness due to the inherent heterogeneity of vegetation growth. Nevertheless, these Fourier-transform-based physical approximations are simple to implement, provide rapid predictions, and have shown great utility in capturing a range of dune phenomena (Durán et al., 2010; Parteli and Herrmann, 2007), making them attractive for incorporation into larger dune evolution models (Walmsley and Howard, 1985; Stam, 1997; Kroy et al., 2002; Duran and Moore, 2013).

2.2 Computational fluid dynamics simulations of flow over dunes

While analytical predictions for the shear stress perturbation have played a major role in morphological studies, CFD approaches have also been important in understanding the flow dynamics over aeolian landforms (Smyth, 2016). One of the earliest is Wippermann and Gross (1986) in which the mesoscale meteorological model FITNAH, a nonhydrostatic, Reynolds-averaged Navier–Stokes (RANS) model, was used in lieu of a theoretical expression as discussed above to simulate the development and migration of a barchan, thus overcoming limitations of the linear predictions. Additionally, numerical simulations have been used to study the effects of varying height and length of an idealized transverse dune (Parsons et al., 2004a, b), varying roughness heights (Wakes et al., 2010), flow separation on the lee side of dunes (Schatz and Herrmann, 2006; Araújo et al., 2013), and airflow over scarps (Hesp and Smyth, 2021; Bauer and Wakes, 2022), as well as blowouts (Smyth et al., 2012), reversing dunes (Jackson et al., 2020), and nebkha foredunes (Hesp and Smyth, 2017; Furtak-Cole et al., 2022). Other example works include Jackson et al. (2011), which compared several turbulence models including RANS $k-\omega$ SST, several hybrid RANS–LES models without surface roughness effects, and an LES–ABL model that included the surface roughness, highlighting the importance of capturing surface roughness effects, as well as Jackson et al. (2013), which presented full 3D RANS simulations for a large dune field. Of particular relevance to the current study, Ferreira et al. (2013) showed that CFD predictions achieve generally good agreement with experimental measurements of shear stress distributions for single and tandem dunes. In general, (OpenFOAM Ltd., 2020) has become a widely used CFD tool for modeling wind dynamics and bed shear stress over topography, including in coastal environments (e.g., Hesp et al., 2015; Jackson et al., 2013).

2.3 Symbolic regression

Existing analytical solutions for the effective bed shear stress distribution over dunes are relatively fast and easy to implement but lack the accuracy and general validity of full CFD simulations. One possible way to bridge this gap is through data-driven techniques. For example, Wakes et al. (2021) trained machine learning models, including random forest and logistic regression approaches, on CFD model output to generate flow predictors. However, there are a number of methods which have yet to be explored or fully evaluated. In particular, symbolic regression (SR) is a long-standing approach that has seen recent growth due to the emergence of new machine learning tools and techniques (Koza, 1994; Champion et al., 2020; Bakarji et al., 2023; Makke and Chawla, 2024). As its name suggests, SR seeks to identify a symbolic expression using a combination of provided input features, symbolic operators, and constants which best captures the relationship between input and output. While SR may lack the accuracy and training efficiency for large, multi-dimensional datasets achieved by deep neural networks, the resulting analytical expressions are more interpretable and often more generalizable than their neural network counterparts.

Some of the most popular methods for discovering the optimal input–output mapping are the Sparse Identification of Nonlinear Dynamics (SINDy) (Brunton et al., 2016), genetic programming-based approaches (Koza, 1994), and more recent efforts to incorporate neural network and deep learning techniques (Sahoo et al., 2018; Petersen et al., 2019). In SINDy and its variants (Rudy et al., 2019; Champion et al., 2019, 2020; Kaheman et al., 2020; Shea et al., 2021; Fasel et al., 2022), a precomputed library of nonlinear terms along with sparse optimization techniques such as the least absolute shrinkage and selection operator (LASSO) (Brunton et al., 2016) or sequential threshold ridge regression (STRidge) (Rudy et al., 2017) are used to determine a sparse, optimal, linear combination of nonlinear terms. While the SINDy approach is computationally efficient, it is known to struggle with noisy data and is limited in the functional forms that it is able to learn (Makke and Chawla, 2024).

On the other hand, genetic programming (GP) approaches utilize data structures, such as expression trees (Koza, 1994) or imperative representations (Brameier and Banzhaf, 2007) that allow for more general functional forms. However, the use of evolutionary strategies to search the space formed by the combination of input features, operators, and possibly real constants incurs a greater computational cost. In general, GP approaches begin with a randomly seeded initial popula-

tion of candidate expressions that is then evolved over several generational cycles. Each cycle consists of evaluating the existing candidates, selecting candidates for mutation, and finally mutating expressions to create a new set of candidates (Koza, 1994). The fitness of all candidates is tracked over the course of many iterations and the best overall candidate is selected as the solution.

One recent GP-based SR tool is PySR (Cranmer, 2023). In addition to the basic GP formulation, it includes regularized evolution for updating population members, a simplify–optimize stage that aims to reduce the number of equivalent but symbolically different expressions and reduce the burden of discovering equations with floating point constants, and a parsimony penalty which aims to maintain equal numbers of candidate expressions at every level of complexity (Cranmer, 2023). In this paper, we apply PySR to CFD simulation data (Cecil et al., 2024) and evaluate its performance compared to the classical KSH functional and a class of reduced-order modeling techniques that we describe next.

2.4 Non-intrusive reduced-order modeling

Another possible way to bridge the gap between classical analytical techniques and full CFD simulations is to build a surrogate (or reduced-order) version of the CFD model itself. If successful, this approach can combine the physical fidelity of a RANS or LES approximation with the speed of a closed-form analytical expression. A vast and growing body of literature exists on reduced-order modeling (e.g., Hesthaven et al., 2016; Benner et al., 2015; Lee and Carlberg, 2020; Brunton et al., 2020). The fundamental idea underlying these approaches is that high-fidelity simulation data often exhibit dynamics with low-dimensional structure. One of the most popular ways to exploit this structure is through the method of snapshots and projection-based model reduction (Carlberg et al., 2017; Benner et al., 2015). In this case, one first generates a set of high-fidelity simulation data (or snapshots) and identifies a low-dimensional approximation to the solution subspace (or manifold) through a dimension reduction step. Reduced-order solutions are then expressed as an expansion in a basis that spans the solution subspace (Dutta et al., 2021a). Traditional techniques like proper orthogonal decomposition (POD) use linear dimension reduction, while nonlinear techniques based on deep neural network architectures are becoming increasingly popular for problems with complex, multi-scale dynamics (Wan et al., 2018; Lee and Carlberg, 2020; Maulik et al., 2021; Dutta et al., 2022).

If the governing equations are known and accessible through the high-fidelity model, one can use Galerkin or Petrov–Galerkin projection to create a rigorous, interpretable reduced-order model (ROM) (Carlberg et al., 2013; Lozovskiy et al., 2016, 2017). Often, this is not practical (or possible) due to design of the high-fidelity model or restrictions on access to its source code. For these cases, purely data-driven methods that do not require intrusive access to the high-fidelity model have become popular (Xiao et al., 2017; Dutta et al., 2021b). Rather than using Galerkin or Petrov–Galerkin projection, these non-intrusive ROMS (NIROMs) use black-box interpolation (or regression) to obtain expansion coefficients for the approximate solution in the reduced basis. A wide array of interpolation techniques and strategies have been tested including Gaussian process regression (Guo and Hesthaven, 2019), dynamic mode decomposition (Tu et al., 2014; Wu et al., 2021), radial basis functions (RBFs) (Walton et al., 2013; Xiao et al., 2015; Dutta et al., 2021a), and neural networks (Hesthaven and Ubbiali, 2018; Salvador et al., 2021; Dutta et al., 2021c). In each case, data for the NIROM interpolation are provided by the projection coefficients of the original simulation snapshots. As a result, the success of the resulting ROM depends heavily on balancing accuracy with over-fitting and having training data that adequately represent the problem space of interest (Dutta et al., 2021a).

Below, we will employ well-established NIROM techniques based on linear dimension reduction and RBF interpolation to build reduced-order approximations of high-fidelity CFD simulations over various dune profiles. These results will serve as benchmarks to evaluate the accuracy and generalizability of our SR approximations.

3 Methodology

3.1 Prescribed dune shapes for modeling

In real-world systems, dunes take on a wide range of shapes and sizes (e.g., Fig. 1). To examine the influence of variable topography on bed shear stress patterns, for this study we examined four types of dune profiles that have been parameterized to allow for consistent specification of the height H and characteristic length L given by the half-length at half-height. These include a quadratic cosine, Gaussian, quartic, and bump profile given by the following.

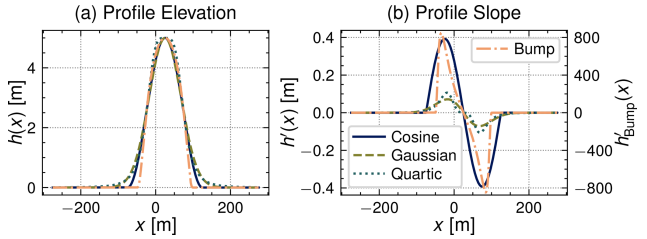

Figure 3. Comparison of prototype dune profile **(a)** and slopes **(b)**. Note that the bump profile slope uses the right y axis in panel **(b)** due to the large difference in scale between it and the other profiles.

$$\text{Cosine: } h(x) = a_c \cos^2[b_c(x-\sigma)]; \quad a_c = H;$$

$$b_c = \frac{1}{L}\arccos\left(\sqrt{\frac{1}{2}}\right); \quad x \in \left[-\frac{\pi}{2b_c}+\sigma, \frac{\pi}{2b_c}+\sigma\right] \tag{6a}$$

$$\text{Gaussian: } h(x) = a_g \exp\left[-\frac{(x-\sigma)^2}{b_g}\right]; \quad a_g = H;$$

$$b_g = -L^2 \ln^{-1}\left(\frac{1}{2}\right) \tag{6b}$$

$$\text{Quartic: } h(x) = \frac{a_q}{1+b_q(x-\sigma)^4}; \quad a_q = H;$$

$$b_q = \frac{1}{L^4} \tag{6c}$$

$$\text{Bump: } h(x) = a_b \exp\left[\frac{-1}{1-b_b(x-\sigma)^2}\right]; \quad a_b = He;$$

$$b_b = \frac{1+[1/\ln(\frac{1}{2})]}{L^2}; x \in \left[\sqrt{\frac{1}{b_b}}+\sigma, \; -\sqrt{\frac{1}{b_b}}+\sigma\right] \tag{6d}$$

Here, σ specifies the x location of the dune peak and h represents the local bed elevation. When applying symbolic regression in Sect. 3.2.3, the shift, σ, becomes important for recovering both real and imaginary parts of the Fourier-transformed expression in Eq. (5).

The profile elevations and slopes are plotted in Fig. 3 for $H/L = 0.1$. It should be noted that while all of the profiles can be controlled to have a small height-to-length ratio as shown in Fig. 3a, only the Gaussian and quartic profiles fulfill the small slope requirement (i.e., $h'(x) = \mathcal{O}(H/L)$) as depicted in Fig. 3b. The cosine profile represents a moderate departure from this constraint and the bump profile a more extreme case that is representative of a scarp. Thus, these profiles allow us to explore not only the effect of increasing H/L ratios on the accuracy of the KSH model, but also the limitations of the underlying small slope assumption when it is not implicitly imposed by a small height-to-length ratio.

For all cases considered, the length-to-aerodynamic-roughness ratio, L/z_0, is held constant at $\Phi = L/z_0 = 25000$ by choosing $L = 25\,\text{m}$ and $z_0 = 1 \times 10^{-3}\,\text{m}$, leading to the KSH model coefficients $A = 5.1952$ and $B = 0.2793$ ($AB = 1.4510$) that remain constant across all cases. Three sets of H/L ratios are considered, namely a small set with $H/L \in [0.01, 0.1]$, medium with $H/L \in [0.1, 0.3]$, and large with $H/L \in [0.3, 0.5]$. In reference back to Fig. 1, the bump profile has $L_{\text{base}} = 37.57\,\text{m}$, leading to a maximum H/L_{base} of 0.3, and for the cosine profile $L_{\text{base}} = 50\,\text{m}$, yielding a maximum H/L_{base} of 0.25. This captures $\sim 35\,\%$ of the coastal foredune aspect ratios presented in Fig. 1 assuming $L_{\text{base}} \approx 2.5\,L$ to account for the quartic and Gaussian profiles. A summary of the case properties is given in Table 1.

3.2 Modeling of shear stress perturbations

3.2.1 CFD simulations

Computational fluid dynamic simulations were run for the medium and large case sets described in the previous section using the open-source solver OpenFOAM. Specifically, the Reynolds-averaged Navier–Stokes (RANS) equations were solved with the RNG $k-\epsilon$ turbulence model using the incompressible SimpleFOAM solver. Second-order accurate discretization schemes were used throughout. More precisely, the least-squares method was applied for determining cell-centered gradient values with a cubic limiter for the velocity components and turbulence quantities. The linear upwind scheme was chosen for divergence terms and the linear scheme for Laplacian terms. At the inlet, an atmospheric boundary layer (ABL) profile (Richards and Hoxey, 1993; Yang et al., 2009) was specified according to

$$u = \frac{u^*}{\kappa}\ln\left(\frac{z-d+z_0}{z_0}\right) \tag{7}$$

$$v = w = 0 \tag{8}$$

$$k = \frac{(u^*)^2}{\sqrt{C_\mu}} \tag{9}$$

$$\epsilon = \frac{(u^*)^3}{\kappa(z-d+z_0)} \tag{10}$$

$$u^* = \frac{u_{\text{ref}}\kappa}{\ln(\frac{z_{\text{ref}}+z_0}{z_0})}, \tag{11}$$

where $u^* = \sqrt{\tau/\rho}$ is the friction velocity, $C_\mu = 0.09$ is a model constant, k the turbulent kinetic energy, and ϵ is the turbulent kinetic energy dissipation rate. In order to maintain the inlet profile over the approach stretch, a constant shear stress condition was applied to the velocity at the top boundary (Richards and Hoxey, 1993; Hargreaves and Wright, 2007; Richards and Norris, 2019). At the surface, a wall function specific to ABL simulations (Richards and Hoxey, 1993; Parente et al., 2011) as implemented in OpenFOAM's `atmEpsilonWallFunction` was used, and the turbulent viscosity was constrained by the `atmNutkWallFunction`. A reference velocity of $u_{\text{ref}} = 10\,\text{m}\,\text{s}^{-1}$ at a reference height of $z_{\text{ref}} = 10\,\text{m}$ was used for all cases.

Table 1. Details of base case sets.

Case set	Profiles	H/L ratios	Φ	A	AB
Small	Cosine, Gaussian, quartic, bump	0.01 – 0.1	25 000	5.1952	1.4510
Medium	Cosine, Gaussian, quartic, bump	0.1 – 0.3	25 000	5.1952	1.4510
Large	Cosine, Gaussian, quartic, bump	0.3 – 0.5	25 000	5.1952	1.4510

The computational domain consisted of a rectangular area that was 425 m long with a total height 200 m greater than the height of the dune crest H. The dunes were positioned such that their crest is located at $x = 30$ m, while the computational domain was symmetric about $x = 0$ m, resulting in the crest being 242.5 m downwind of the inlet and 182.5 m from the outflow boundary. A uniform base discretization of 1 m was created using blockMesh and subsequently refined near the surface to between 0.0625 and 0.125 m using SnappyHexMesh. Additionally, approximately 20 anisotropic surface layers were inserted with the first cell height adjusted per case to maintain y^+ values between approximately 30 and 200. This corresponds to a minimum first cell height of 5.138×10^{-3} m obtained for the quartic profile with $H/L = 0.5$. Simulations were run to steady state with convergence established when all residuals had fallen to at least 1×10^{-8} and the iteration-to-iteration change of the calculated drag over the hill profile was less than 1×10^{-8}.

3.2.2 CFD data preparation

The geometry representations and shear stress results of the CFD simulations provide a dataset to which data-driven model discovery tools can be applied. To begin, the surface topography and total shear stress were extracted using OpenFOAM's sampling utilities. Even spacing between samples for subsequent Fourier transforms was ensured by resampling the data at an equivalent number of points using cubic spline interpolation. Furthermore, some of the leading and trailing ends of the domain (i.e., 150 points, ~ 9 m, at either end of the 6800 total samples, ~ 425 m) were truncated to avoid spurious boundary affects. The bed shear stress perturbation was then recovered using Eq. (1) where the value at the leading edge of the truncated domain was used for normalization. Furthermore, the results were filtered to remove high-frequency noise in both the topography and shear stress and downsampled to $N = 650$ samples to obtain a reasonably sized, high-fidelity dataset for performing SR and reduced-order modeling.

3.2.3 Symbolic regression

Using the processed CFD data, discrete Fourier transforms (DFTs) were taken of the topography and shear stress perturbation. These along with the Fourier wave numbers were provided to PySR to solve the minimization problem,

$$\min_{f()} \left|\left| \mathcal{F}\left[\tau'\right] - f\left(\mathcal{F}\left[h(x)\right], k\right) \right|\right|_2, \tag{12}$$

in complex Fourier space where $f()$ denotes an unknown function residing in the space of operators consisting of addition, subtraction, multiplication, and absolute value as well as additional complex constants. PySR's regularized evolution algorithm was run for 100 iterations where each iteration consists of multiple rounds of tournament selection and mutation (Cranmer, 2023). Several of the default hyperparameters were adjusted in order to obtain consistent results across all cases for several different levels of synthetic noise (see Sect. 4.2.1). A listing of non-default options that were supplied to PySR is available in Table A1 located in Appendix A. Finally, after obtaining the optimal expression $f(\cdot)$ from PySR, the bed shear stress perturbation in real space is given by

$$\tau' = \mathcal{F}^{-1}\left[f\left(\mathcal{F}\left[h(x)\right], k\right)\right], \tag{13}$$

where $\mathcal{F}^{-1}[\cdot]$ indicates the inverse Fourier transform.

3.2.4 Non-intrusive reduced-order modeling

In addition to the CFD and symbolic regression approaches, we also explore predictions based on linear NIROM techniques. As discussed earlier, this basic approach relies on selecting a low-dimensional, data-driven basis for the approximate solution that ideally captures essential solution behavior at much reduced computational cost. That is, we write the approximate solution $\tilde{\boldsymbol{y}}$ in $\mathbb{R}^N$ on the downsampled CFD grid as

$$\tilde{\boldsymbol{y}} = \sum_{j=1}^{m} z_j \boldsymbol{u}_j, \tag{14}$$

where $\{\boldsymbol{u}_j\}_{j=1}^m$, $\boldsymbol{u}_j \in \mathbb{R}^N$ is the basis and ideally $m \ll N$. Next, a regression model is constructed to capture the parameterized relationship between the input topography and output surface shear stress perturbation. To do this, we represent the topography, $\boldsymbol{h} \in \mathbb{R}^N$, on the downsampled CFD grid as

$$\boldsymbol{h} = \sum_{j=1}^{m} c_j \boldsymbol{d}_j, \tag{15}$$

where $\{\boldsymbol{d}_j\}_{j=1}^m$, $\boldsymbol{d}_j \in \mathbb{R}^N$ is a low-dimensional basis for the topography space. A regression metamodel $G : \boldsymbol{c} \rightarrow \boldsymbol{z}$ then

maps the discrete input to output coefficients $\{z_j\}$, and the approximate solution is given by Eq. (14).

For an alternative basis, we explore either a truncated Fourier basis in which the original theoretical model resides or a POD expansion (Rathinam and Petzold, 2003). There are pros and cons to each reduced space approximation. POD provides an optimal representation of the training (snapshot) data in terms of the Euclidean 2-norm (Quarteroni et al., 2016) and hence better data compression, while a Fourier basis requires more modes for smooth reconstruction (≈ 80 modes for the simulations considered below) but provides better generalization to unseen data.

Below, for the POD basis we combine both the input topography profiles and resulting CFD-based shear stress perturbations into a composite dataset and generate a single basis for both. That is, we perform n_s training simulations and compute a singular value decomposition:

$$\mathbf{S} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{W}^T, \tag{16}$$

where the columns of the snapshot matrix $\mathbf{S} \in \mathbb{R}^{N\times 2n_s}$ hold both the input topographies $\{\boldsymbol{d}_j\}_{j=1}^{n_s}$ and shear stress perturbations $\{\boldsymbol{y}_j\}_{j=1}^{n_s}$. The first m columns of $\mathbf{U}$, $\{\boldsymbol{u}_j\}_{j=1}^{m}$ form the solution and topography basis (i.e., $\boldsymbol{d}_j = \boldsymbol{u}_j \; \forall j = 1, \ldots, m$). Then for a given topography input evaluated on the CFD grid, $\boldsymbol{h}$, we compute the projection $\boldsymbol{c} = \mathbf{D}^T\boldsymbol{h}$, where $\mathbf{D} \in \mathbb{R}^{N\times m}$ is a matrix holding the basis elements $\{\boldsymbol{d}_j\}$ as columns and similarly for the shear stress perturbation to obtain $\boldsymbol{z}$.

For the Fourier basis, $\boldsymbol{d}_j$ and $\boldsymbol{u}_j$ are obtained through DFTs, which are truncated by zeroing elements beyond the selected number of modes. Subsequently, the complex Fourier representations are split into real and imaginary parts and individual metamodels constructed for $\mathcal{R}(G) : \mathcal{R}(\boldsymbol{c}) \rightarrow \mathcal{R}(\boldsymbol{z})$ and $\mathcal{I}(G) : \mathcal{I}(\boldsymbol{c}) \rightarrow \mathcal{I}(\boldsymbol{z})$. The final predictions are then generated via

$$\tilde{\boldsymbol{y}} = \mathcal{F}^{-1}\big[\mathcal{R}(G) + \mathcal{I}(G)\big]. \tag{17}$$

For constructing the regression metamodels in the latent space we use RBF interpolation to capture the input–output mapping with either a Gaussian kernel using a shape parameter of 0.1 and additional linear polynomial (for the POD basis) or a multi-quadric kernel using a shape parameter equal to 0.1 and a zero-degree polynomial (for the Fourier basis). While we proceed here by choosing either a random set across all four profiles or all cases of two profiles to construct the POD basis and train the metamodel, there are alternative selection criteria that could be considered such as the standard and greedy approaches found in Dutta et al. (2021a). Furthermore, there is still significant flexibility within the above framework for how the input parameter space is sampled to generate snapshots and how these are used to generate the reduced bases and regression mappings. For example, a separate POD basis could be constructed for the input and output or regression performed in complex space for the Fourier-based approach.

3.3 Model skill metrics

To assess differences between the various models, we examine the mean square error (MSE) defined by

$$\mathrm{MSE} = (1/N)\sum_{i=0}^{N}(y_i - \tilde{y}_i)^2, \tag{18}$$

where the CFD results are taken as the target values, y_i, and the KSH, SR, or NIROM results as the predictions, $\tilde{y}_i$, as well as the maximum error across all spatial locations. In the results presented below all errors are calculated for the shear stress perturbations and are therefore analogous to a relative error for the total bed shear stress.

4 Results

4.1 CFD predictions

We begin by examining the results of the CFD simulations compared to the theoretical predictions of Eq. (5) in Fig. 4. Representative results for the surface shear stress distribution are provided for the bump and Gaussian profiles for two H/L ratios in Fig. 4a–d. Our CFD simulations predict separation in the lee for $H/L \geq 0.3$ for the bump profile, $H/L \geq 0.38$ for the quartic, and $H/L \geq 0.5$ for the cosine, while the flow remains attached for all of the Gaussian cases considered.

In what follows, it should be noted that we have compared our CFD results to unmodified predictions from the other models. In practice, the shear stress perturbation prediction from KSH is usually limited to values greater than or equal to -1 in recognition of validity constraints imposed by the underlying assumptions. On the lee side, separation streamlines are typically defined using phenomenological fits of either a cubic polynomial (Kroy et al., 2002) or ellipse (Schatz and Herrmann, 2006) based on the dune geometry to improve prediction on the stoss slope. However, these approaches prove difficult to implement when a clear brink or separation location cannot be identified, as in the case of the current profiles, and fail to account for additional factors controlling the extent of the separating streamline such as the upwind shear velocity (Walker and Nickling, 2002; Araújo et al., 2013). For these reasons, and in an effort to produce a consistent comparison among all models, we have used the unmodified predictions of τ' over the actual idealized profiles.

The MSE and maximum error are plotted for each of the four profiles and 21 H/L ratios in Fig. 4e and f. Here we can clearly see that the error increases with increasing H/L as is to be expected from the assumptions present in the KSH model. Additionally, the errors for the bump profile, which drastically violates the low-slope assumption while satisfying a small H/L, are greater than the other profiles for each value of H/L. In fact, referring back to Fig. 3b, the MSE increases for increasing max slope except for the quartic

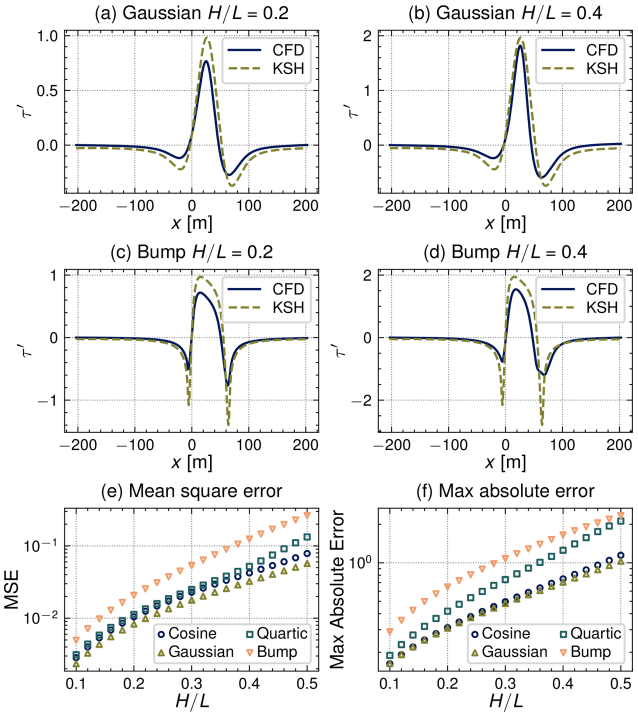

Figure 4. Comparison of CFD simulation results and KSH predictions. Bed shear stress perturbation predictions for selected cases are presented in panels **(a–d)**. The mean square error for each case is given panel **(e)** and the maximum absolute error in panel **(f)**.

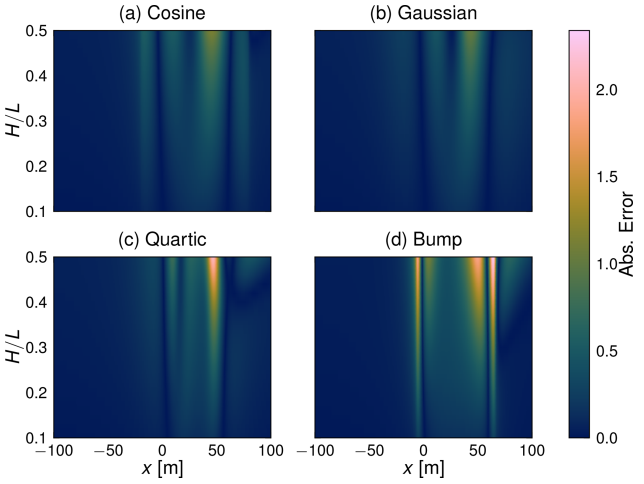

Figure 5. Spatial distribution of absolute difference between OpenFOAM and KSH bed shear stress perturbation predictions. Note that the x domain has been limited to highlight the area of interest immediately surrounding the dune.

case, which is roughly equivalent to the cosine profile until $H/L \approx 0.4$, at which point the MSE for the quartic profile begins to increase at a faster rate.

Next, we examine the spatial distribution of absolute error in Fig. 5. Interestingly, the error in maximum shear stress prediction (just ahead of the dune crests situated at $x = 30$ m) from the KSH model either stays relatively constant or decreases for increasing H/L. The most drastic increase in prediction error comes on the lee side of the dune profiles where separation is most likely to occur. Again, considering the slope of the different profiles, the bump profile exhibits two distinct regions of higher error just fore and aft of the main dune body where the slopes are highest. This effect is drastically reduced in the other profiles which exhibit much gentler slope characteristics.

Finally, the differences in the surface shear stress perturbation extrema predicted by CFD and KSH are tabulated in Table 2. The difference in the local minima near the dune toe consistently increases with increasing H/L. However, the differences in maxima near the crest and local minima near the heel exhibit more nuanced behavior depending on the profile. For example, the cosine, Gaussian, and quartic profile difference magnitudes in maximum τ' increase until $H/L \approx 0.26$ before decreasing as H/L continues to increase. However, for the bump profile, the absolute value of the difference in maximum τ' increases consistently for the range of H/L considered. A similar observation can be made for the minima in the lee, except for the cosine profile, which exhibits a consistent increase in difference as well. Considering this in light of Figs. 5 and 4e, which show a consistent increase in absolute error and MSE, respectively, suggests that the overall errors are due not only to differences in magnitude but also differences in the phase prediction with respect to the topography and overall distribution of τ'.

4.2 Symbolic regression predictions

4.2.1 Symbolic regression on KSH

To begin exploring the use of symbolic regression via PySR, we first applied the methodology to data directly generated by the KSH model. After generating shear stress perturbation predictions for each of the hill profiles and height-to-length ratios using Eq. (5), we applied Gaussian noise at several signal-to-noise ratios (SNRs) to the shear stress perturbations. This was done in order to evaluate PySR's robustness to noise in measured output data in anticipation of dealing with the potentially noisy CFD predictions. PySR was trained on each of the small, medium, and large case sets using 8 cases for training and the remaining 36 cases for validation.

The average MSEs for each of the sets and SNRs considered are depicted in Fig. 6. In nearly all cases, PySR was able to learn the KSH model form and coefficient values. For those cases in which agreement of the symbolic form was not exact, e.g., the large set with SNR = 20, the resulting expression still exhibited very low error. A table of each of the discovered expressions is supplied in Appendix A. Having determined that the PySR symbolic regression package was capable of discovering the underlying expression for analyti-

Table 2. Differences in surface shear stress extrema predictions between CFD and KSH. Values are reported for local minima near the dune toe and heel, as well as the maximum near the dune crest for each profile and H/L considered.

	Cosine			Gaussian			Quartic			Bump		
H/L	Toe	Crest	Heel	Toe	Crest	Heel	Toe	Crest	Heel	Toe	Crest	Heel
0.100	0.088	−0.144	0.092	0.052	−0.143	0.052	0.066	−0.141	0.060	0.249	−0.141	0.269
0.120	0.107	−0.165	0.112	0.062	−0.163	0.061	0.079	−0.160	0.071	0.308	−0.165	0.336
0.140	0.126	−0.184	0.131	0.072	−0.180	0.071	0.092	−0.176	0.080	0.369	−0.189	0.406
0.160	0.146	−0.200	0.150	0.083	−0.195	0.080	0.104	−0.190	0.089	0.434	−0.211	0.479
0.180	0.166	−0.213	0.170	0.093	−0.206	0.090	0.117	−0.201	0.096	0.502	−0.232	0.554
0.200	0.186	−0.225	0.191	0.103	−0.215	0.099	0.130	−0.208	0.102	0.572	−0.253	0.631
0.220	0.207	−0.233	0.211	0.113	−0.221	0.108	0.142	−0.213	0.107	0.645	−0.271	0.712
0.240	0.228	−0.239	0.231	0.123	−0.224	0.117	0.155	−0.217	0.109	0.720	−0.289	0.796
0.260	0.249	−0.243	0.251	0.133	−0.224	0.125	0.167	−0.218	0.109	0.798	−0.306	0.885
0.280	0.271	−0.244	0.269	0.143	−0.221	0.133	0.179	−0.217	0.105	0.877	−0.321	0.979
0.300	0.293	−0.243	0.287	0.153	−0.216	0.140	0.192	−0.215	0.099	0.959	−0.336	1.079
0.320	0.316	−0.239	0.303	0.164	−0.207	0.146	0.205	−0.212	0.089	1.042	−0.350	1.179
0.340	0.339	−0.232	0.317	0.173	−0.196	0.151	0.217	−0.205	0.077	1.128	−0.363	1.278
0.360	0.362	−0.225	0.330	0.184	−0.183	0.153	0.230	−0.198	0.064	1.214	−0.376	1.381
0.380	0.386	−0.214	0.340	0.194	−0.167	0.154	0.242	−0.190	0.050	1.302	−0.388	1.483
0.400	0.410	−0.202	0.348	0.204	−0.148	0.152	0.255	−0.184	0.035	1.392	−0.400	1.584
0.420	0.435	−0.189	0.353	0.214	−0.127	0.147	0.268	−0.178	0.017	1.482	−0.411	1.683
0.440	0.460	−0.174	0.357	0.224	−0.104	0.139	0.280	−0.175	0.013	1.574	−0.423	1.785
0.460	0.485	−0.158	0.360	0.234	−0.078	0.127	0.293	−0.174	0.011	1.667	−0.434	1.898
0.480	0.511	−0.143	0.362	0.245	−0.051	0.111	0.306	−0.176	0.000	1.761	−0.446	2.015
0.500	0.537	−0.128	0.363	0.255	−0.023	0.092	0.318	−0.177	0.005	1.856	−0.458	2.128

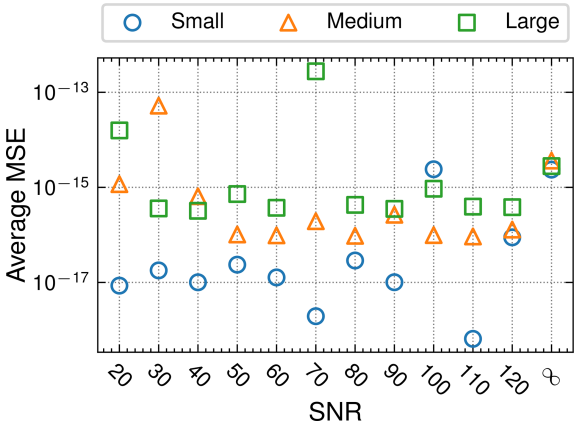

Figure 6. Average MSE across all profiles for each set of H/L ratios in Table 1 and several SNRs.

cally generated data in the Fourier domain and was robust to noise in the output variable, we turned our attention to applying SR to the CFD-generated data.

4.2.2 Symbolic regression on CFD

Having shown that PySR is capable of discovering the KSH model from directly generated data, we next applied PySR to the processed CFD results for the medium set of H/L values for which the flow remains attached for all cases. For this set, PySR returned a symbolic expression that nearly matches the form of the KSH model, differing only in the value of the multiplicative constants with AB picking up a negligible real part given by

$$\mathcal{F}[\tau'] = \left(3.29|k| + \left(1.62i - 1.8 \times 10^{-6}\right)k\right)\mathcal{F}[h]. \tag{19}$$

It should be noted that physics-based constraints were not explicitly enforced in the SR training that led to Eq. (19). However, physics considerations are implicitly included since the training data were provided by physics-based CFD models and we have performed the regression on the Fourier transform of the input–output features in following with the asymptotic solutions. We believe this emphasizes the applicability of the linearized, asymptotic predictions to cases where separation is not present as well as the potential for SR to discover generalized expressions from data reflecting the underlying physics in some scenarios.

Being of the same form as KSH, Eq. (19) essentially represents an optimization of the model coefficients. The discovered value for $A = 3.29$ is significantly lower than suggested by KSH, while $B = 0.492$ is greater. It is known that the asymptotic expressions in Eq. (5) overpredict the value of A while underpredicting the value of AB (Charru et al., 2013). The results presented in Charru et al. (2013) obtained using a full numerical solution to the turbulent boundary layer predict A between 3 and 4 depending on the closure model, while AB is less than 2, which is in agreement with our re-

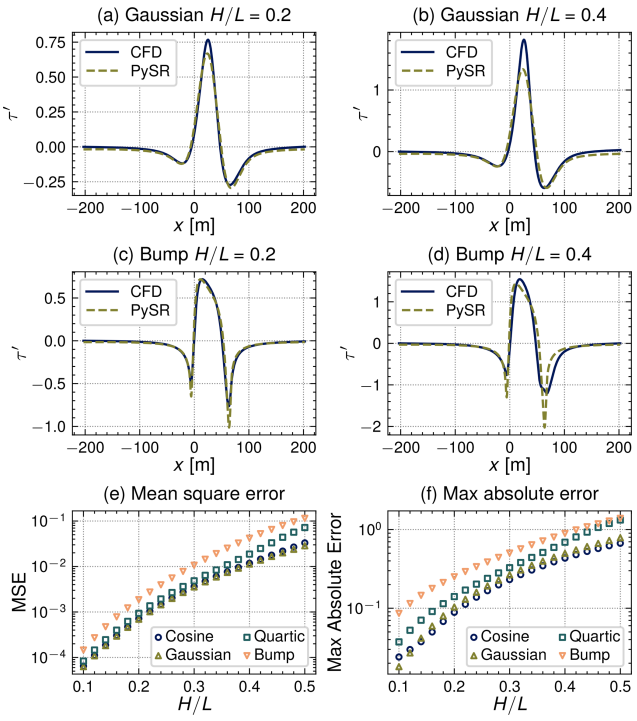

Figure 7. Comparison of CFD and symbolic regression predictions on the medium case set, i.e., Eq. (19). Selected bed shear stress perturbation profiles are shown in panels **(a–d)** with the MSE and max absolute error in panels **(e)** and **(f)**, respectively.

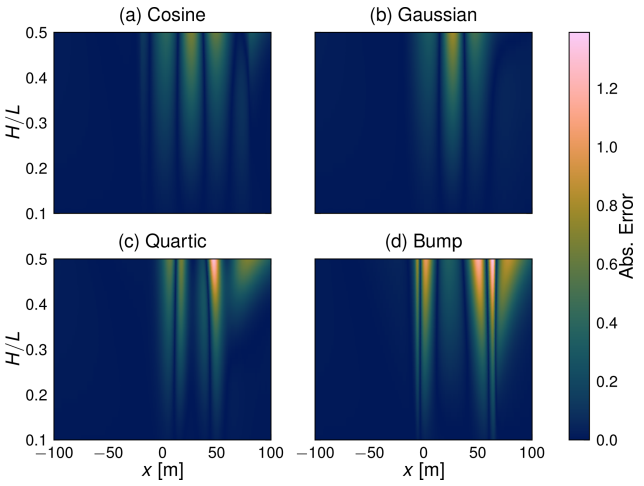

Figure 8. Spatial absolute difference between CFD and symbolic regression trained on select cases from the medium case set, i.e., Eq. (19).

sult of $A = 3.29$ and $AB = 1.62$. It should also be noted that the error between the full solution and asymptotic predictions for A decreases as L/z_0 increases, while the error for AB remains relatively constant (Charru et al., 2013). This suggests that our CFD results, and thus SR predictions for A and B, would be closer to the KSH predictions for higher values of L/z_0.

Results using Eq. (19) are summarized in Fig. 7. As can be seen in Fig. 7a–d, the PySR model tends to underpredict the maximum shear stress perturbation in contrast to the standard KSH predictions while still overpredicting the minima just fore and aft of the dune. However, overall the PySR discovered model results in roughly an order of magnitude improvement in MSE for H/L values between 0.1 and 0.3 as shown in Fig. 7e. However, results are more muted when extending this model to the larger H/L values for which separation may occur. This can largely be attributed to the limitations of the linearized KSH model, and thus the SR learned model with an identical functional form, to be able to capture separation effects without ad hoc corrections. Additionally, when considering the spatial distribution of absolute error in Fig. 8, the error near the dune crest increases with successively higher values of H/L, in contrast to results for KSH in Fig. 5. Also, the shape of the shear stress perturbation near the peak of the bump profile is better captured, while the largest error still occurs just fore and aft where the slopes are steepest.

Differences in the τ' extrema predictions between CFD and Eq. (19) are tabulated in Table 3. The differences in local minima near the dune toe are consistently improved over their KSH counterparts. However, the differences in maximum τ' are only improved for low H/L and exhibit a monotonic increase for increasing H/L, showing degraded predictions compared to KSH for H/L greater than approximately 0.28, except for the bump profile which exhibits consistent improvement over KSH. In the lee, results are again more nuanced depending on the profile considered. Improvement over KSH is seen for the cosine, Gaussian, and bump profiles, while the results for the quartic are somewhat degraded. Again, considering these results along with Fig. 8, it appears that Eq. (19) captures the phase advance and overall distribution of τ' more accurately (particularly for the cosine, Gaussian, and bump profiles) than KSH for the cases considered, and hence the overall MSE can largely be attributed to differences in magnitude.

In an effort to obtain better generalizability to separated cases, we next applied PySR to the large case set with somewhat unsatisfying results. Several trials were run using PySR while varying the number of cases provided, the cases selected for training, i.e., separated only or separated and attached scenarios, and the operator space to search. The most promising result obtained takes the form of the KSH model with several additional terms that are quadratic in the wave number given by

$$\mathcal{F}[\tau'] = \big(-2.32\mathcal{F}[h]|k| + \mathcal{F}[h]\big)\big[\big(3.26 \times 10^{-6} + i2.20\big)k + 4.84|k| - 0.0380 + i1.81 \times 10^{-7}\big]. \quad (20)$$

Note that Eq. (20) contains terms representing the second derivative of the dune profile, thus increasing the model sensitivity to the underlying topography representation. Results obtained using Eq. (20) are summarized in Fig. 9. The trends

Table 3. Differences in surface shear stress extrema predictions between CFD and Eq. (19). Values are reported for local minima near the dune toe and heel, as well as the maximum near the dune crest for each profile and H/L considered.

	Cosine			Gaussian			Quartic			Bump		
H/L	Toe	Crest	Heel	Toe	Crest	Heel	Toe	Crest	Heel	Toe	Crest	Heel
0.100	0.006	0.009	0.021	0.002	0.014	0.013	−0.002	0.015	0.011	0.035	−0.014	0.078
0.120	0.008	0.019	0.026	0.002	0.025	0.015	−0.003	0.028	0.011	0.051	−0.013	0.106
0.140	0.011	0.031	0.031	0.003	0.039	0.017	−0.003	0.043	0.010	0.070	−0.011	0.138
0.160	0.014	0.045	0.036	0.003	0.056	0.019	−0.004	0.060	0.009	0.092	−0.008	0.172
0.180	0.017	0.062	0.042	0.003	0.076	0.021	−0.006	0.081	0.007	0.117	−0.004	0.209
0.200	0.021	0.082	0.048	0.003	0.099	0.022	−0.007	0.104	0.003	0.145	0.001	0.248
0.220	0.025	0.104	0.054	0.004	0.124	0.023	−0.008	0.130	−0.003	0.175	0.008	0.290
0.240	0.029	0.129	0.060	0.004	0.153	0.025	−0.009	0.158	−0.010	0.208	0.015	0.336
0.260	0.034	0.156	0.065	0.004	0.183	0.026	−0.010	0.188	−0.021	0.242	0.024	0.388
0.280	0.040	0.185	0.070	0.004	0.218	0.026	−0.011	0.220	−0.034	0.279	0.034	0.443
0.300	0.045	0.217	0.073	0.004	0.255	0.025	−0.012	0.253	−0.051	0.318	0.044	0.504
0.320	0.051	0.251	0.075	0.004	0.295	0.023	−0.013	0.288	−0.070	0.359	0.056	0.566
0.340	0.058	0.289	0.075	0.004	0.337	0.020	−0.014	0.326	−0.092	0.401	0.068	0.627
0.360	0.065	0.327	0.073	0.004	0.382	0.015	−0.015	0.365	−0.116	0.445	0.081	0.691
0.380	0.072	0.368	0.069	0.004	0.429	0.008	−0.016	0.403	−0.139	0.491	0.094	0.755
0.400	0.080	0.411	0.062	0.004	0.479	−0.002	−0.017	0.441	−0.165	0.537	0.107	0.818
0.420	0.088	0.454	0.054	0.005	0.532	−0.014	−0.018	0.478	−0.192	0.585	0.121	0.879
0.440	0.096	0.499	0.044	0.005	0.586	−0.030	−0.019	0.513	−0.206	0.634	0.135	0.942
0.460	0.105	0.546	0.031	0.005	0.643	−0.049	−0.020	0.544	−0.218	0.685	0.149	1.017
0.480	0.114	0.593	0.020	0.005	0.702	−0.073	−0.021	0.573	−0.239	0.736	0.163	1.095
0.500	0.124	0.639	0.007	0.006	0.761	−0.100	−0.022	0.604	−0.244	0.788	0.176	1.170

in Fig. 9a–d are similar to the previous results; however, for the bump profile, while the minimums appear to be predicted better, some oscillatory behavior also appears in the results. This could be attributed to the inclusion of second derivative terms and the large spikes in the second derivative of the bump profile in those areas. Concerning the max error in Fig. 9f, the bump and cosine profiles exhibit a slower increase in maximum error until $H/L \approx 0.35$. The spatial distribution of error in Fig. 10 also exhibits trends similar to the previous results.

The differences in τ' extrema predictions between CFD and Eq. (20) are tabulated in Table 4. Results are generally comparable to the results obtained using Eq. (19) in Table 3, especially for the local minima near the toe and the maximum near the crest. For the minima in the lee, results are more profile-dependent, with a decrease in accuracy seen for the cosine and Gaussian cases, but improvements are gained for the quartic and bump profiles. Overall, Eq. (20) provides some modest improvement for errors over the bump profile, while little is gained for the other profiles considered. This could be attributed to the fact that the majority of the separated cases were obtained for the bump profile, thus biasing the symbolic regression model. A more balanced future dataset should include an expanded H/L range to accommodate additional separated flow cases for other profiles.

4.3 NIROM predictions

Due to difficulties with accurately fitting separated cases using PySR, an alternative approach was explored which turned to non-intrusive reduced-order modeling techniques as discussed in Sect. 3.2.4. Results for these methods are illustrated in Fig. 11. The DFT–RBF results are presented in Fig. 11c and e and the POD–RBF in Fig. 11d and f. The Fourier basis was limited to 80 modes after which changes in the reconstruction were not visible for additional modes. For the POD basis, 14 modes were retained for the scattered training selection based on retaining 99.9 % of the energy, while only 8 modes were required for the two-profile training selection. The normalized singular values for both training selections are presented in Fig. 11a and b.

For the scattered training selection, a random sampling of 40 cases across all four profiles was chosen as the training set. The training cases were used to build the POD basis as well as train the RBF interpolator for both DFT–RBF and POD–RBF. Results for this approach are given in Fig. 11c and d. As to be expected, the highest errors occur for points farthest from neighboring cases of the same profile and for extrapolation at the edges of the case domain. Furthermore, while the DFT–RBF approach achieves better reconstruction errors on average, the POD–RBF results offer a slight improvement in predicting unseen values of H/L over the DFT–RBF.

Table 4. Differences in surface shear stress extrema predictions between CFD and Eq. (20). Values are reported for local minima near the dune toe and heel, as well as the maximum near the dune crest for each profile and H/L considered.

	Cosine			Gaussian			Quartic			Bump		
H/L	Toe	Crest	Heel	Toe	Crest	Heel	Toe	Crest	Heel	Toe	Crest	Heel
0.100	0.019	0.002	0.068	0.022	0.013	0.063	0.020	0.039	0.057	−0.026	−0.029	0.025
0.120	0.024	0.010	0.083	0.026	0.025	0.075	0.024	0.056	0.067	−0.022	−0.031	0.042
0.140	0.029	0.021	0.097	0.030	0.039	0.087	0.027	0.076	0.075	−0.015	−0.033	0.064
0.160	0.035	0.034	0.112	0.035	0.055	0.098	0.031	0.098	0.083	−0.005	−0.033	0.088
0.180	0.041	0.050	0.127	0.039	0.075	0.110	0.034	0.123	0.090	0.008	−0.032	0.114
0.200	0.047	0.068	0.142	0.043	0.097	0.121	0.038	0.151	0.095	0.024	−0.029	0.142
0.220	0.054	0.089	0.157	0.047	0.123	0.132	0.041	0.182	0.100	0.041	−0.025	0.173
0.240	0.061	0.112	0.172	0.051	0.151	0.144	0.044	0.215	0.101	0.062	−0.021	0.209
0.260	0.069	0.138	0.187	0.055	0.182	0.155	0.047	0.249	0.100	0.084	−0.015	0.250
0.280	0.076	0.166	0.201	0.059	0.216	0.165	0.051	0.285	0.096	0.109	−0.009	0.295
0.300	0.085	0.196	0.214	0.063	0.253	0.174	0.054	0.324	0.088	0.136	−0.001	0.345
0.320	0.093	0.229	0.225	0.068	0.293	0.182	0.057	0.364	0.078	0.165	0.008	0.396
0.340	0.103	0.265	0.234	0.072	0.335	0.188	0.061	0.405	0.066	0.195	0.017	0.446
0.360	0.112	0.302	0.242	0.076	0.380	0.194	0.064	0.449	0.051	0.227	0.026	0.500
0.380	0.122	0.342	0.247	0.080	0.427	0.197	0.067	0.492	0.037	0.260	0.037	0.554
0.400	0.132	0.383	0.250	0.084	0.477	0.196	0.071	0.536	0.021	0.295	0.047	0.606
0.420	0.143	0.426	0.250	0.088	0.528	0.194	0.074	0.577	0.002	0.331	0.058	0.657
0.440	0.154	0.468	0.250	0.092	0.581	0.188	0.078	0.616	−0.003	0.368	0.069	0.709
0.460	0.166	0.513	0.247	0.097	0.640	0.178	0.081	0.652	−0.005	0.405	0.079	0.773
0.480	0.178	0.560	0.245	0.100	0.699	0.165	0.085	0.684	−0.017	0.445	0.089	0.842
0.500	0.190	0.603	0.242	0.105	0.759	0.148	0.088	0.722	−0.013	0.485	0.101	0.905

The second training selection criteria consisted of using all cases from two profiles as training data while predicting on the unseen profiles (see Fig. 11e and f). This allowed us to explore the generalizability of the NIROM methods to unseen topography shapes rather than just unseen parameters as in the first scenario. For this experiment, the DFT–RBF results in Fig. 11e far outperform the POD–RBF results in Fig. 11f. This is most likely due to the less universal nature of the POD basis, which is constructed purely based on the training data and is unable to accurately represent both the input topography and shear stress perturbation for completely unseen profiles. On the other hand, since the DFT basis is not dependent on the data being examined, it is capable of representing the unseen topography and the error is due to extrapolating the input–output relationship to unseen profiles.

5 Discussion

5.1 Intercomparison of shear stress predictions

Thus far we have analyzed the differences in surface shear stress perturbation predictions between an established analytical model, namely KSH, CFD simulations, and data-driven approaches. We have shown that KSH generally overpredicts the extrema compared to an RNG $k-\epsilon$ RANS formulation for a range of two-dimensional dune profiles. Errors increase with increasing H/L and have been shown to be correlated with topography slope regardless of the overall dune dimensions. Furthermore, we have explored two data-driven methodologies as avenues for improving computationally efficient models compared to resource-intensive CFD simulations. These various simulations and approaches were compared through the lens of assessing existing capabilities and alternative options for resolving spatial wind and bed shear stress calculations within existing process-based aeolian sediment transport models.

Having examined each of the shear stress perturbation prediction models individually, all of the results are gathered in Fig. 12 in order to discuss their relative performance. It can be seen that KSH is outperformed by all of the trained models for the range of H/L ratios and profiles considered, except for the POD–RBF 2×2 results. This result is not too surprising since each of the alternatives was optimized for this specific dataset during training. Concerning the SR results, the expression trained only on the medium set of H/L values, i.e., PySR Med, outperforms that trained only on large H/L values, i.e., PySR Large; however, the prediction error becomes more similar as H/L increases, with PySR Large gaining a slight advantage for the bump and quartic profiles. Overall, the SR results for the medium case set confirm the validity of linearized, asymptotic representations such as KSH while representing a modest improvement due to an optimized prediction of A and B in agreement with Charru et al. (2013). The impact of these modest differences on morphological predictions is discussed further in the following sec-

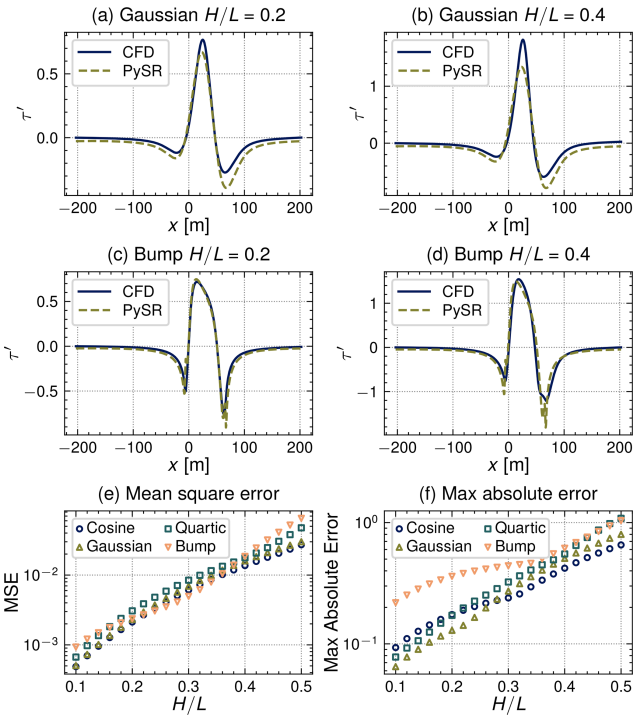

Figure 9. Comparison of CFD and symbolic regression predictions trained on the large case set, i.e., Eq. (20). Selected bed shear stress perturbation profiles are provided in panels **(a–d)** with the MSE and max absolute error for all cases in panels **(e)** and **(f)**, respectively.

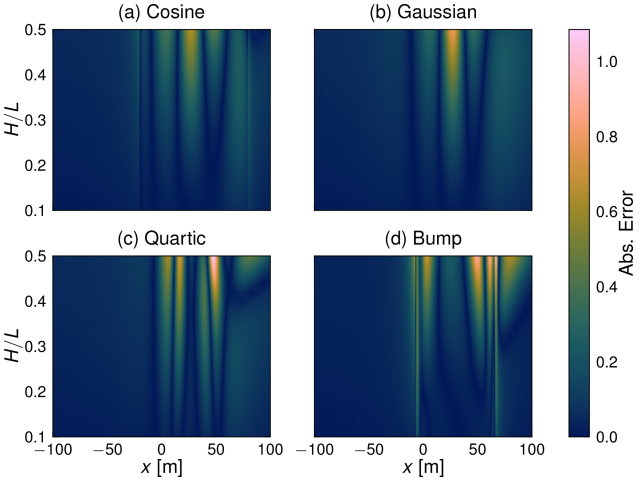

Figure 10. Spatial absolute difference between CFD and symbolic regression trained on cases from the large case set, i.e., Eq. (20).

tion. However, as demonstrated by the PySR Large results, work is still needed for reducing the error for larger values of H/L and for capturing the distribution of surface shear stress in regions of flow separation.

Linear, projection-based, non-intrusive, reduced-order modeling approaches were able to further reduce the prediction error at the cost of generalizability to unseen cases. For

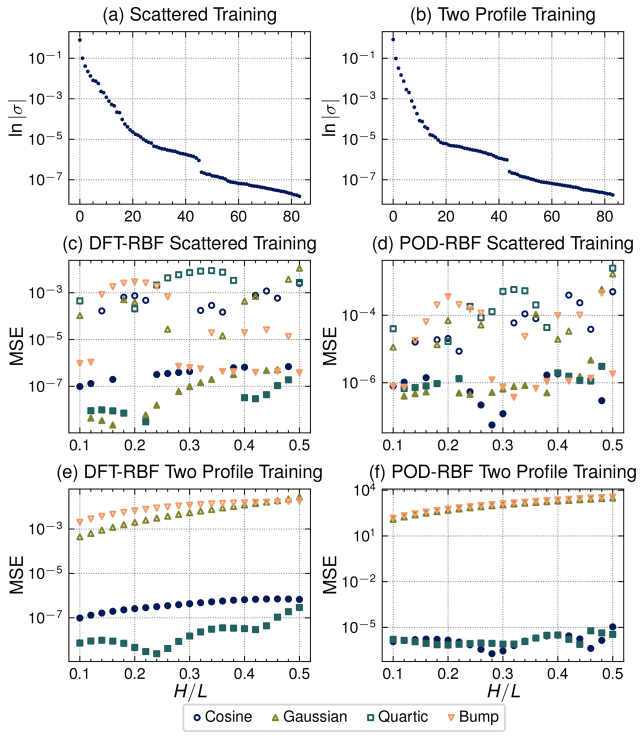

Figure 11. Summary of results for the NIROM-based shear stress perturbation prediction models. Panels **(a)** and **(b)** depict the normalized singular values for the training data used to construct the POD basis. Panels **(c)** and **(d)** show the MSE for the DFT–RBF and POD–RBF models when using randomly selected cases as training data. Panels **(e)** and **(f)** present the MSE when choosing all cases from two profiles as training data. In panels **(c–f)** filled marks indicate cases used for training where the reported MSE represents the reconstruction error; empty marks represent test cases where the reported MSE represents the prediction error compared to the CFD results.

the NIROM approaches, scattered training, which covers the entire data space, performs best. Prediction errors are largely dependent on the distance from training cases, and in the case of POD–RBF 2×2, the model is unable to generalize to unseen profiles, which poses a challenge for real-world applications where topography evolves with time. Compared to the SR results, the prediction errors are no worse and exhibit significant improvement for high H/L values, especially when unseen cases fall close to seen cases as is true for the bump profile. The success of the simple NIROM approaches in this region, and in particular the DFT–RBF variant, suggests that data-driven methodologies may be able to provide fast prediction models which accurately capture separation effects.

5.2 Error implications for morphological prediction

For theoretical applications of aeolian landform evolution, the exact magnitude of transport rates matters less than capturing the patterns that lead to sediment transport gradients

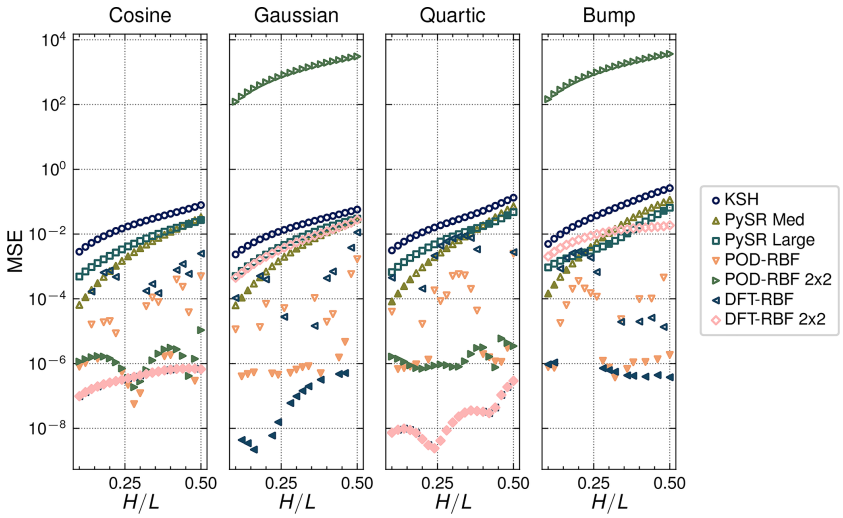

Figure 12. MSE for all shear stress perturbation models grouped by profile. The POD–RBF 2×2 and DFT–RBF 2×2 results refer to the two profile-trained models. Filled marks indicate cases used for training where the reported MSE represents the reconstruction error; empty marks represent test cases where the reported MSE represents the prediction error compared to the CFD results.

that spur net landform changes. Numerical tools, largely using the KSH approximation, have yielded the ability to synthesize complex two- or three-dimensional topographic change patterns and behaviors as observed in nature (e.g., Durán et al., 2010; Kombiadou et al., 2023). Yet, limited field-scale validation exists in terms of these behaviors. Increasingly, there is a push for more engineering-oriented capabilities that rely on appropriately simulating relevant physical phenomena that necessitate accurate prediction of spatiotemporal rates of sediment mobilization and transport. Ultimately, sediment transport rates are driven by excess shear above a threshold to initiate grain motion. Numerous formulations for aeolian sediment transport exist in the literature, many but not all of which (Martin and Kok, 2017) express the sediment flux (q) as a nonlinear function of u_* or τ. For example, the commonly used modified Bagnold equation (Bagnold, 1937) defines sediment flux as

$$q = C \frac{\rho_a}{g} \sqrt{\frac{d}{D}} (u_* - u_{*,t})^3, \tag{21}$$

where ρ_a is the air density, g is gravity, d is the grain size, and D is a reference grain size. $u_{*,t}$ is the threshold shear velocity for initial grain motion, which may be estimated as

$$u_{*,t} = A \sqrt{\frac{\rho_g - \rho_a}{\rho_a} g d}, \tag{22}$$

where A is a coefficient related to sediment distribution spread and ρ_g is the grain density.

Given that τ and corresponding u_* profiles over dune topography are highly nonuniform resulting from flow acceleration, deceleration, and separation bubble effects and that deviations in KSH from the CFD model results are often not insignificant, this has important implications for the resulting potential for aeolian transport across the dune face.

To directly assess rates of transport and corresponding bed elevation change patterns, we use KSH and PySR outputs as input hot-start files to the AeoLiS (Hoonhout and de Vries, 2016) model for select simulations. On the upwind end of the domain, CFD and PySR outputs both indicate biased higher τ relative to KSH at the base of the dune (e.g., Fig. 13e, f). As spatial gradients in sediment transport result in net bed elevation change, this local reduction in shear results in a small zone of deposition. For the $H = 6\,\text{m}$, $L = 25\,\text{m}$ case for $\tau = 0.6\,\text{Pa}$, KSH results in about a 75 % larger deposition height at the base of a Gaussian-shaped dune. As previously noted, deviations between KSH and PySR are generally larger on the dune itself, with maximum differences occurring at the crest of the dune form. Assuming bare sand conditions (i.e., no vegetation), spatial gradients in τ predicted by both KSH and PySR result in increasing q across the dune face and therefore a net landward migration of the landform. However, assuming the CFD-based PySR results better represent the true wind flow dynamics, Fig. 14 shows the excess net landward sediment transport across the dune crest from the two spatial shear stress models. Notably, as the wind conditions become more energetic or as the H/L ratio increases, errors associated with the magnitude of sand transport increase. For a given boundary τ, these H/L versus KSH–PySR model differences follow an exponential relationship (Fig. 14).

Taken in sum, mis-characterization of the shear distribution across the dune profile has important implications for accurately simulating coastal landform change in terms of both the shape and rate of change. CFD is generally too computationally demanding to couple with aeolian transport mod-

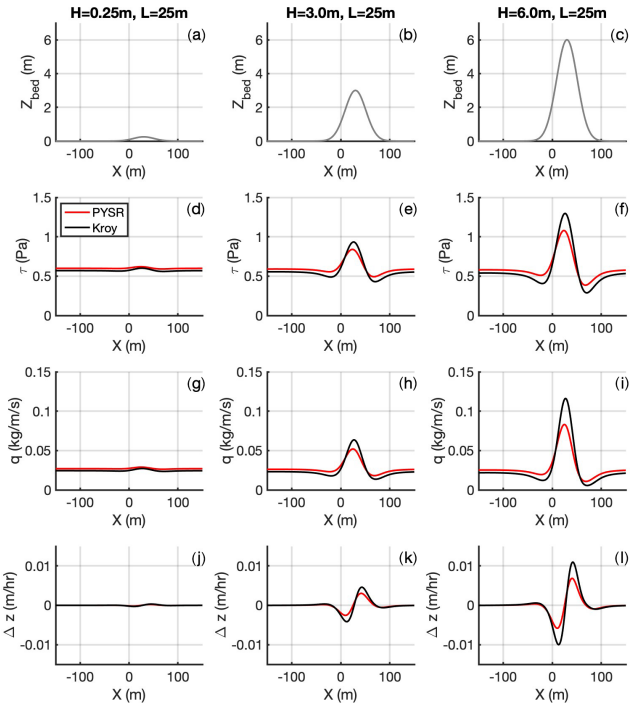

Figure 13. Example differences in predicted aeolian sediment transport fields **(g–i)** and bed elevation change **(j–l)** from KSH and PySR for $\tau = 0.6$ Pa for H/L ratios of 0.01 **(a, d, g, j)**, 0.12 **(b, e, h, k)**, and 0.24 **(c, f, i, l)** Gaussian-shaped profiles.

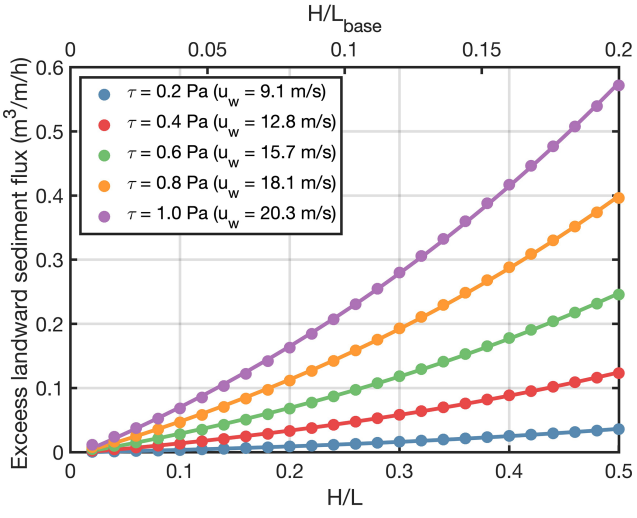

Figure 14. Net differences in simulated sediment flux past the dune crest between KSH and PySR for various H/L conditions (Gaussian-shaped dune) and boundary τ. Dots represent simulated AeoLiS cases and lines reflect the best-fit exponential curve through data points. For linkage to Fig. 1, $L_{\text{base}} = 2.5L$ is assumed.

els, requiring surrogate or simplified modeling. Given that SR shows the ability to recreate the KSH form for simplified geometries and improve upon flow errors, SR-type approaches could potentially play an important role for improving process-based numerical modeling tools for landform evolution that are required for engineering and design needs. Such approaches are particularly important given that $\sim 99\,\%$, $93\,\%$, and $63\,\%$ of coastal dunes from the database in Fig. 1 have H/L ratios exceeding 0.1, 0.2, and 0.5, respectively, assuming $L_{\text{base}} = 2.5\,L$. However, these findings are likely to also have broad implications for landform modeling in other arid (e.g., Tsoar, 2001; Parteli et al., 2014) or extraterrestrial settings (e.g., Atwood-Stone and McEwen, 2013; Parteli and Herrmann, 2007) where slope angles can similarly locally approach the angle of repose.

5.3 Synthesis and remaining challenges

Symbolic regression was able to discover an expression similar to KSH that offers an order of magnitude improvement in prediction error when compared to CFD; however, it still suffers from many of the same issues as KSH, namely, that the error increases with increasing H/L, and is unable to capture the effects of separation. Linear, projection-based, non-intrusive, reduced-order modeling approaches were able to further reduce the prediction error at the cost of generalizability to unseen cases. While we were able to obtain modest prediction improvements, we recognize that several challenges remain.

First, in this study we only considered a single value of L/z_0 which controls the value of the model constants in KSH and would most likely play a role in any alternative discovered model. Future efforts should extend the dataset to include a range of L/z_0 in order to try to extract a relation governing the SR-discovered model coefficients. Only then could the expression be incorporated into existing coastal dune evolution models in a general sense.

Second, we have not considered the effects of vegetation within this study. Steeper coastal foredunes are often vegetated, resulting in reduced shear at the sand surface. This somewhat reduces the concern of errors for steep dunes as the surface shear often drops below the transport threshold; however, in many models the effect of vegetation is applied as a correction factor to the topographically driven shear stress perturbation as in Durán and Moore (2013). Therefore, seeking to increase the accuracy of predictions of τ' over the equivalent unvegetated surface for steep dunes is still warranted.

Third, while the NIROM approaches show promise, especially if data are available that sufficiently cover the range of dune topographies expected for a given scenario, concern for their ability to generalize to unexpected topographies which may arise over the course of a dune evolution simulation remains. However, other techniques, such as the use of nonlinear auto-encoders to determine the reduced basis representation and deep learning methods for capturing the latent space mapping from input to output, may provide viable alternatives with improved generalization (Maulik et al., 2021; Dutta et al., 2022).

Finally, we have focused exclusively on simple, idealized two-dimensional dune profiles and have not considered more complex transects that may experience multiple areas of separation such as can occur with severely scarped profiles. Furthermore, real-world coastal dunes are inherently three-dimensional and experience a variety of wind directions and complex, three-dimensional flow effects such as topographic steering and recirculation (e.g., Smyth et al., 2012; Piscioneri et al., 2019). A not insignificant amount of work would be required to extend these approaches to three-dimensional dunes considering the increased computational expense for generating 3D CFD simulation data, as well as considering topographic variations in an additional dimension. However, the methodologies themselves should be directly applicable given data availability.

6 Conclusions

Coastal dunes represent an important line of defense against storm-related flooding hazards. As such, accurate models for predicting their evolution over relevant engineering timescales is paramount for future risk assessments and the development of engineering solutions. Current process-based models make use of linearized analytical solutions to the underlying governing equations which, while easy to implement and computationally efficient, have known deficits for steep topographies. This work highlights that the KSH form widely used in existing tools performs well for the low-slope dune shapes for which the analytical formulation is designed. On dune forms with steeper topography and higher H/L ratios, errors in the magnitude of bed shear stress predictions – compared to CFD model predictions – increase across the dune face. These errors are typically largest at the dune crest, the ramification of which is that existing formulations result in a faster net landward migration of the dune form in morphological models than would be expected. Overall, we have provided an analysis of the limitations of existing analytical prediction models for airflow over coastal dunes and presented two data-driven approaches for constructing alternative fast prediction methods that reduce error relative to the KSH approach. The various newly presented approaches reduce the global error of stress predictions and/or improve on the ability to simulate patterns in complex flow dynamics not fully resolved by the analytical solution. While the improvements presented do not yet provide a complete solution for flow predictions across all dune forms, this work both highlights existing limitations in current dune modeling frameworks and demonstrates new methodologies that could enable more reliable predictions of spatial bed shear stress patterns in the full range of H/L conditions observed in natural coastal systems.

Appendix A: Symbolic regression details

This appendix provides additional details for symbolic regression results. PySR hyperparameters that were changed from their default values are provided in Table A1. These values were tuned to provide robust results across all of the noisy KSH-generated data and used in subsequent PySR runs.

An exhaustive list of the expressions returned by PySR when provided KSH-generated data with Gaussian noise is provided in Table A2. In some cases, some basic simplification or regrouping of the expressions has been performed.

Table A1. Robust parameters for PySR.

Parameter name	Description	Value
niterations		100
ncyclesperiteration		300
populations		40
maxsize		20
population_size		500
fraction_replaced		0.0001
fraction_replaced_hof		0.0001
annealing		false
batching		true
binary_operators		$+, \times, -$
unary_operators		$abs()$

Table A2. Simplified PySR discovered expressions for various SNRs. In the interest of space $\hat{}$ symbols indicate Fourier-transformed values.

SNR	Small	Medium	Large
20	$\hat{h}(1.451ik + 5.195\|k\|)$	$\hat{h}(1.451ik + 5.195\|k\|)$	$(1.451ik + \|k - 2k\,(3.087 - 0.236i)\,\|)$
30	$\hat{h}(1.451ik + 5.195\|k\|)$	$\hat{h}[k(-1.087 \times 10^{-6} + 1.451i) + 5.195\|k\|]$	$\hat{h}(1.451ik + 5.195\|k\|)$
40	$\hat{h}(1.451ki + 5.195\|k\|)$	$\hat{h}(1.451ik + 5.195\|k\|)$	$\hat{h}(1.451ik + 5.195\|k\|)$
50	$\hat{h}\left[k(1.0 + 1.451i) + \|k - 5.195\|k\|\|\right]$	$\hat{h}(1.451ik + 5.195\|k\|)$	$\hat{h}(1.451ik + 5.195\|k\|)$
60	$\hat{h}(1.451ik + 5.195\|k\|)$	$\hat{h}(1.451ik + 5.195\|k\|)$	$\hat{h}(1.451ik + 5.195\|k\|)$
70	$2\hat{h}(0.7255ik + 2.598\|k\|)$	$\hat{h}(1.451ik + 5.195\|k\|)$	$\hat{h}\left[1.451ik + \left(5.195 + 4.326 \times 10^{-6}i\right)\|k\|\right]$
80	$\hat{h}(1.451ik + 5.195\|k\|)$	$\hat{h}(1.451ik + 5.195\|k\|)$	$\hat{h}(1.451ik + 5.195\|k\|)$
90	$5.195\hat{h}(0.2793ik + \|k\|)$	$1.451i\hat{h}k + 5.195\hat{h}\|k\|$	$\hat{h}(1.451ik + 5.195\|k\|)$
100	$\hat{h}(1.451ik + 5.195\|k\|)$	$\hat{h}(1.451ik + 5.195\|k\|)$	$\hat{h}(1.451ik + 5.195\|k\|)$
110	$\hat{h}(1.451ik + 5.195\|k\|)$	$\hat{h}(1.451ik + 5.195\|k\|)$	$1.451i\hat{h}k + 5.195\hat{h}\|k\|$
120	$\hat{h}(1.451ik + 5.195\|k\|)$	$5.195\hat{h}(0.2793ik + \|k\|)$	$\hat{h}(1.451ik + 5.195\|k\|)$
∞	$\hat{h}(1.451ik + 5.195\|k\|)$	$\hat{h}(k(1 \times 10^{-6} + 1.451i) + 5.195\|k\|)$	$\hat{h}(1.451ik + 5.195\|k\|)$

Code availability. OpenFOAM v2006 was used for all CFD simulations as distributed by OpenCFD Ltd. Final PySR (Cranmer, 2023) results were generated using version 0.14.3 available at https://github.com/MilesCranmer/PySR (last access: 27 July 2023) (Cranmer, 2023). Version 2.1 of the AeoLiS model (de Vries et al., 2023), used for the aeolian sediment transport calculations in this work, is preserved at https://doi.org/10.4121/22215562, available via GPL-3.0, and developed openly at https://github.com/openearth/aeolis-python (last access: 17 December 2024). Figures 2–11 were prepared using the Batlow color scheme provided by Scientific Color Maps (https://doi.org/10.5281/zenodo.1243862, Crameri, 2023).

Data availability. The processed CFD results (Cecil et al., 2024) used to produce the figures and train the symbolic regression and non-intrusive reduced-order models in this study are available at https://doi.org/10.17603/ds2-4w1m-7998.

Author contributions. The study was conceptualized by NC, OC, MF, and SD. CFD simulations, symbolic regression, and non-intrusive reduced-order modeling were performed by OC with input from MF, SD, and NC. Analysis of the impact on sediment fluxes was performed by NC. OC and NC prepared the original draft with contributions from MF and SD. All authors contributed to review and editing. NC and AT were responsible for funding acquisition.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Acknowledgements. Permission was granted by the Chief of Engineers to publish this information.

Financial support. This research has been partially supported by the ERDC Basic Research Program (project "Wind-Driven Processes in Dryland Environments") and NOAA's National Centers for Coastal Ocean Science Competitive Research Program (grant no. NA19NOS4780175).

Review statement. This paper was edited by Andreas Baas and reviewed by Orencio Duran Vinent and one anonymous referee.

References

Araújo, A. D., Parteli, E. J. R., Pöschel, T., Andrade, J. S., and Herrmann, H. J.: Numerical modeling of the wind flow over a transverse dune, Sci. Rep., 3, 2858, https://doi.org/10.1038/srep02858, 2013.

Atwood-Stone, C. and McEwen, A. S.: Avalanche slope angles in low-gravity environments from active Martian sand dunes, Geophys. Res. Lett., 40, 2929–2934, https://doi.org/10.1002/grl.50586, 2013.

Bagnold, R. A.: The transport of sand by wind, The Geographical Journal, 89, 409–438, https://doi.org/10.2307/1786411, 1937.

Bakarji, J., Champion, K., Nathan Kutz, J., and Brunton, S. L.: Discovering governing equations from partial measurements with deep delay autoencoders, P. R. Soc. A, 479, 20230422, https://doi.org/10.1098/rspa.2023.0422, 2023.

Bauer, B. O. and Wakes, S. J.: CFD simulations of wind flow across scarped foredunes: Implications for sediment pathways

and beach–dune recovery after storms, Earth Surf. Proc. Land., 47, 2989–3015, https://doi.org/10.1002/esp.5439, 2022.

Benner, P., Gugercin, S., and Willcox, K.: A survey of projection-based model reduction methods for parametric dynamical systems, SIAM Rev., 57, 483–531, https://doi.org/10.1137/130932715, 2015.

Brameier, M. F. and Banzhaf, W.: Linear Genetic Programming, Springer US, ISBN 9780387310299, https://doi.org/10.1007/978-0-387-31030-5, 2007.

Bruno, L., Horvat, M., and Raffaele, L.: Windblown sand along railway infrastructures: A review of challenges and mitigation measures, J. Wind Eng. Ind. Aerod., 177, 340–365, https://doi.org/10.1016/j.jweia.2018.04.021, 2018.

Brunton, S. L., Proctor, J. L., and Kutz, J. N.: Discovering governing equations from data by sparse identification of nonlinear dynamical systems, P. Natl. Acad. Sci. USA, 113, 3932–3937, https://doi.org/10.1073/pnas.1517384113, 2016.

Brunton, S. L., Noack, B. R., and Koumoutsakos, P.: Machine learning for fluid mechanics, Annu. Rev. Fluid Mech., 52, 477–508, https://doi.org/10.1146/annurev-fluid-010719-060214, 2020.

Carlberg, K., Farhat, C., Cortial, J., and Amsallem, D.: The GNAT method for nonlinear model reduction: Effective implementation and application to computational fluid dynamics and turbulent flows, J. Comput. Phys., 242, 623–647, https://doi.org/10.1016/j.jcp.2013.02.028, 2013.

Carlberg, K., Barone, M., and Antil, H.: Galerkin v. least-squares Petrov–Galerkin projection in nonlinear model reduction, J. Comput. Phys., 330, 693–734, https://doi.org/10.1016/j.jcp.2016.10.033, 2017.

Cecil, O., Cohn, N., Farthing, M., Dutta, S., and Trautz, A.: 2D OpenFOAM simulations of flow over several idealistic dune profiles using the RNG k-epsilon turbulence model, DesignSafe-CI [data set], https://doi.org/10.17603/ds2-4w1m-7998, 2024.

Champion, K., Lusch, B., Kutz, J. N., and Brunton, S. L.: Data-driven discovery of coordinates and governing equations, P. Natl. Acad. Sci. USA, 116, 22445–22451, https://doi.org/10.1073/pnas.1906995116, 2019.

Champion, K., Zheng, P., Aravkin, A. Y., Brunton, S. L., and Kutz, J. N.: A unified sparse optimization framework to learn parsimonious physics-informed models from data, IEEE Access, 8, 169259–169271, https://doi.org/10.1109/access.2020.3023625, 2020.

Charbonneau, B. R., Duarte, A., Swannack, T. M., Johnson, B. D., and Piercy, C. D.: DOONIES: A process-based ecogeomorphological functional community model for coastal dune vegetation and landscape dynamics, Geomorphology, 398, 108037, https://doi.org/10.1016/j.geomorph.2021.108037, 2022.

Charru, F., Andreotti, B., and Claudin, P.: Sand Ripples and Dunes, Annu. Rev. Fluid Mech., 45, 469–493, https://doi.org/10.1146/annurev-fluid-011212-140806, 2013.

Cohn, N.: Dune Response Tool, GitHub [data set], https://github.com/erdc/dune-response-tool (last access: 17 December 2024), 2022.

Crameri, F.: Scientific colour maps, Zenodo [code], https://doi.org/10.5281/zenodo.1243862, 2023.

Cranmer, M.: Interpretable machine learning for science with PySR and SymbolicRegression.jl, arXiv [preprint], https://doi.org/10.48550/ARXIV.2305.01582, 2023 (code available at: https://github.com/MilesCranmer/PySR).

de Vries, S., Hallin, C., van IJzendoorn, C., van Westen, B., Cohn, N., Strypsteen, G., Skaden, J., Agrawal, N., and Garcia Alvarez, M.: AeoLiS, Zenodo [code], https://doi.org/10.4121/22215562, 2023.

Dickey, J., Wengrove, M., Cohn, N., Ruggiero, P., and Hacker, S. D.: Observations and modeling of shear stress reduction and sediment flux within sparse dune grass canopies on managed coastal dunes, Earth Surf. Proc. Land., 48, 907–922, 2023.

Diniega, S., Glasner, K., and Byrne, S.: Long-time evolution of models of aeolian sand dune fields: Influence of dune formation and collision, Geomorphology, 121, 55–68, https://doi.org/10.1016/j.geomorph.2009.02.010, 2010.

Doran, K., Long, J., Birchler, J., Brenner, O., Hardy, M., Morgan, K., Stockdon, H., and Torres, M.: Lidar-derived beach morphology (dune crest, dune toe, and shoreline) for U.S. sandy coastlines, Tech. rep., https://doi.org/10.5066/F7GF0S0Z, 2017.

Duran, O. and Moore, L. J.: Vegetation controls on the maximum size of coastal dunes, P. Natl. Acad. Sci. USA, 110, 17217–17222, https://doi.org/10.1073/pnas.1307580110, 2013.

Durán, O. and Moore, L. J.: Vegetation controls on the maximum size of coastal dunes, P. Natl. Acad. Sci. USA, 110, 17217–17222, 2013.

Durán, O., Parteli, E. J., and Herrmann, H. J.: A continuous model for sand dunes: Review, new developments and application to barchan dunes and barchan dune fields, Earth Surf. Proc. Land., 35, 1591–1600, https://doi.org/10.1002/esp.2070, 2010.

Durán Vinent, O. and Moore, L. J.: Barrier island bistability induced by biophysical interactions, Nat. Clim. Change, 5, 158–162, https://doi.org/10.1038/nclimate2474, 2015.

Dutta, S., Farthing, M. W., Perracchione, E., Savant, G., and Putti, M.: A greedy non-intrusive reduced order model for shallow water equations, J. Comput. Phys., 439, 110378, https://doi.org/10.1016/j.jcp.2021.110378, 2021a.

Dutta, S., Rivera-Casillas, P., Cecil, O., Farthing, M., Perracchione, E., and Putti, M.: Data-driven reduced order modeling of environmental hydrodynamics using deep autoencoders and neural ODEs, in: 9th edition of the International Conference on Computational Methods for Coupled Problems in Science and Engineering, online, June 2021, CIMNE, https://doi.org/10.23967/coupled.2021.017, 2021b.

Dutta, S., Rivera-Casillas, P., and Farthing, M. W.: Neural ordinary differential equations for data-driven reduced order modeling of environmental hydrodynamics, in: Proceedings of the AAAI 2021 Spring Symposium on Combining Artificial Intelligence and Machine Learning with Physical Sciences, Standford University, Palo Alto, CA, March 2021, CEUR-WS, https://doi.org/10.48550/arXiv.2104.13962, 2021c.

Dutta, S., Rivera-Casillas, P., Styles, B., and Farthing, M. W.: Reduced Order Modeling Using Advection-Aware Autoencoders, Mathematical and Computational Applications, 27, 34, https://doi.org/10.3390/mca27030034, 2022.

Elko, N., Brodie, K., Stockdon, H., Nrodstrom, K., Houser, C., McKenna, K., Moore, L., Rosati, J., Ruggiero, P., Thuman, R., and Walker, I.: Dune management challenges on developed coasts, Shore Beach, 84, 15–28, 2016.

Fasel, U., Kutz, J. N., Brunton, B. W., and Brunton, S. L.: Ensemble-SINDy: robust sparse model discovery in the low-data, high-noise limit, with active learning and control, P. R.

Soc. A, 478, 20210904, https://doi.org/10.1098/rspa.2021.0904, 2022.

Ferreira, A. D., Pinheiro, S. R., and Francisco, S. C.: Experimental and numerical study on the shear velocity distribution along one or two dunes in tandem, Environ. Fluid Mech., 13, 557–570, https://doi.org/10.1007/s10652-013-9282-7, 2013.

Figlus, J.: Designing and implementing coastal dunes for flood risk reduction, in: Coastal Flood Risk Reduction, Elsevier, 287–301, https://doi.org/10.1016/B978-0-323-85251-7.00021-4, 2022.

Furtak-Cole, E., Gillies, J. A., Hilgendorf, Z., Walker, I. J., and Nikolich, G.: Simulation of flow and shear stress distribution on the Oceano Dunes, implications for saltation and dust emissions, Environ. Fluid Mech., 22, 1399–1420, https://doi.org/10.1007/s10652-022-09902-0, 2022.

Guo, M. and Hesthaven, J.: Data-driven reduced order modeling for time-dependent problems, Comput. Method. Appl. M., 345, 75–99, https://doi.org/10.1016/j.cma.2018.10.029, 2019.

Hage, P., Ruessink, G., Van Aartrijk, Z., and Donker, J.: Using video monitoring to test a fetch-based aeolian sand transport model, Journal of Marine Science and Engineering, 8, 110, https://doi.org/10.3390/jmse8020110, 2020.

Hallin, C., van IJzendoorn, C., Homberger, J.-M., and de Vries, S.: Simulating surface soil moisture on sandy beaches, Coast. Eng., 185, 104376, https://doi.org/10.1016/j.coastaleng.2023.104376, 2023.

Hanley, M., Hoggart, S., Simmonds, D., Bichot, A., Colangelo, M., Bozzeda, F., Heurtefeux, H., Ondiviela, B., Ostrowski, R., Recio, M., Trude, R., Zawadzka-Kahlau, E., and Thompson, R.: Shifting sands? Coastal protection by sand banks, beaches and dunes, Coast. Eng., 87, 136–146, https://doi.org/10.1016/j.coastaleng.2013.10.020, 2014.

Hargreaves, D. and Wright, N.: On the use of the $k-\epsilon$ model in commercial CFD software to model the neutral atmospheric boundary layer, J. Wind Eng. Ind. Aerod., 95, 355–369, https://doi.org/10.1016/j.jweia.2006.08.002, 2007.

Hesp, P. A. and Smyth, T. A.: Nebkha flow dynamics and shadow dune formation, Geomorphology, 282, 27–38, https://doi.org/10.1016/j.geomorph.2016.12.026, 2017.

Hesp, P. A. and Smyth, T. A.: CFD flow dynamics over model scarps and slopes, Phys. Geogr., 42, 1–24, https://doi.org/10.1080/02723646.2019.1706215, 2021.

Hesp, P. A., Smyth, T. A., Nielsen, P., Walker, I. J., Bauer, B. O., and Davidson-Arnott, R.: Flow deflection over a foredune, Geomorphology, 230, 64–74, https://doi.org/10.1016/j.geomorph.2014.11.005, 2015.

Hesthaven, J. S. and Ubbiali, S.: Non-intrusive reduced order modeling of nonlinear problems using neural networks, J. Comput. Phys., 363, 55–78, https://doi.org/10.1016/j.jcp.2018.02.037, 2018.

Hesthaven, J. S., Rozza, G., and Stamm, B.: Certified Reduced Basis Methods for Parametrized Partial Differential Equations, Springer, https://doi.org/10.1007/978-3-319-22470-1, 2016.

Hoonhout, B. M. and de Vries, S.: A process-based model for aeolian sediment transport and spatiotemporal varying sediment availability, J. Geophys. Res.-Earth, 121, 1555–1575, https://doi.org/10.1002/2015JF003692, 2016.

Hunt, J. C. R., Leibovich, S., and Richards, K. J.: Turbulent shear flows over low hills, Q. J. Roy. Meteor. Soc., 114, 1435–1470, https://doi.org/10.1002/qj.49711448405, 1988.

Jackson, D., Cruz-Avero, N., Smyth, T., and Hernández-Calvento, L.: 3D airflow modelling and dune migration patterns in an arid coastal dune field, J. Coastal Res., 165, 1301–1306, https://doi.org/10.2112/si65-220.1, 2013.

Jackson, D. W., Cooper, A., Green, A., Beyers, M., Guisado-Pintado, E., Wiles, E., Benallack, K., and Balme, M.: Reversing transverse dunes: Modelling of airflow switching using 3D computational fluid dynamics, Earth Planet. Sc. Lett., 544, 116363, https://doi.org/10.1016/j.epsl.2020.116363, 2020.

Jackson, D. W. T., Beyers, J. H. M., Lynch, K., Cooper, J. A. G., Baas, A. C. W., and Delgado-Fernandez, I.: Investigation of three-dimensional wind flow behaviour over coastal dune morphology under offshore winds using computational fluid dynamics (CFD) and ultrasonic anemometry, Earth Surf. Proc. Land., 36, 1113–1124, https://doi.org/10.1002/esp.2139, 2011.

Jackson, P. S. and Hunt, J. C. R.: Turbulent wind flow over a low hill, Q. J. Roy. Meteor. Soc., 101, 929–955, https://doi.org/10.1002/qj.49710143015, 1975.

OpenFOAM Ltd.: OpenFOAM v2006 User Guide Documentation, https://openfoam.com/documentation/guides/v2006/doc/ (last access: 17 December 2024), 2020.

Kaheman, K., Kutz, J. N., and Brunton, S. L.: SINDy-PI: A robust algorithm for parallel implicit sparse identification of nonlinear dynamics, P. R. Soc. A, 476, 20200279, https://doi.org/10.1098/rspa.2020.0279, 2020.

Keijsers, J., De Groot, A., and Riksen, M.: Modeling the biogeomorphic evolution of coastal dunes in response to climate change, J. Geophys. Res.-Earth, 121, 1161–1181, https://doi.org/10.1002/2015JF003815, 2016.

Khalaf, F. and Al-Ajmi, D.: Aeolian processes and sand encroachment problems in Kuwait, Geomorphology, 6, 111–134, https://doi.org/10.1016/0169-555X(93)90042-Z, 1993.

Kombiadou, K., Costas, S., and Roelvink, D.: Exploring controls on coastal dune growth through a simplified model, J. Geophys. Res.-Earth, 128, e2023JF007080, https://doi.org/10.1029/2023JF007080, 2023.

Koza, J.: Genetic programming as a means for programming computers by natural selection, Stat. Comput., 4, 87–112, https://doi.org/10.1007/bf00175355, 1994.

Kroy, K., Sauermann, G., and Herrmann, H. J.: Minimal model for aeolian sand dunes, Phys. Rev. E, 66, 031302, https://doi.org/10.1103/physreve.66.031302, 2002.

Lee, K. and Carlberg, K. T.: Model reduction of dynamical systems on nonlinear manifolds using deep convolutional autoencoders, J. Comput. Phys., 404, 108973, https://doi.org/10.1016/j.jcp.2019.108973, 2020.

Lozovskiy, A., Farthing, M., Kees, C., and Gildin, E.: POD-based model reduction for stabilized finite element approximations of shallow water flows, J. Comput. Appl. Math., 302, 50–70, https://doi.org/10.1016/j.cam.2016.01.029, 2016.

Lozovskiy, A., Farthing, M., and Kees, C.: Evaluation of Galerkin and Petrov–Galerkin model reduction for finite element approximations of the shallow water equations, Comput. Method. Appl. M., 318, 537–571, https://doi.org/10.1016/j.cma.2017.01.027, 2017.

Luna, M. C. d. M., Parteli, E. J., Durán, O., and Herrmann, H. J.: Model for the genesis of coastal dune fields with vegetation, Geomorphology, 129, 215–224, https://doi.org/10.1016/j.geomorph.2011.01.024, 2011.

Makke, N. and Chawla, S.: Interpretable scientific discovery with symbolic regression: a review, Artif. Intell. Rev., 57, 2, https://doi.org/10.1007/s10462-023-10622-0, 2024.

Martin, R. L. and Kok, J. F.: Wind-invariant saltation heights imply linear scaling of aeolian saltation flux with shear stress, Science Advances, 3, e1602569, https://doi.org/10.1126/sciadv.1602569, 2017.

Mason, P. J. and Sykes, R. I.: Flow over an isolated hill of moderate slope, Q. J. Roy. Meteor. Soc., 105, 383–395, https://doi.org/10.1002/qj.49710544405, 1979.

Maulik, R., Lusch, B., and Balaprakash, P.: Reduced-order modeling of advection-dominated systems with recurrent neural networks and convolutional autoencoders, Phys. Fluids, 33, 037106, https://doi.org/10.1063/5.0039986, 2021.

Moore, L. J., Vinent, O. D., and Ruggiero, P.: Vegetation control allows autocyclic formation of multiple dunes on prograding coasts, Geology, 44, 559–562, https://doi.org/10.1130/G37778.1, 2016.

Mull, J. and Ruggiero, P.: Estimating storm-induced dune erosion and overtopping along U.S. west coast beaches, J. Coastal Res., 298, 1173–1187, https://doi.org/10.2112/jcoastres-d-13-00178.1, 2014.

Okin, G. S.: A new model of wind erosion in the presence of vegetation, J. Geophys. Res., 113, F02S10, https://doi.org/10.1029/2007jf000758, 2008.

Parente, A., Gorlé, C., van Beeck, J., and Benocci, C.: Improved $k-\epsilon$ model and wall function formulation for the RANS simulation of ABL flows, J. Wind Eng. Ind. Aerod., 99, 267–278, https://doi.org/10.1016/j.jweia.2010.12.017, 2011.

Parsons, D., Wiggs, G., Walker, I., Ferguson, R., and Garvey, B.: Numerical modelling of airflow over an idealised transverse dune, Environ. Modell. Softw., 19, 153–162, https://doi.org/10.1016/s1364-8152(03)00117-8, 2004a.

Parsons, D. R., Walker, I. J., and Wiggs, G. F.: Numerical modelling of flow structures over idealized transverse aeolian dunes of varying geometry, Geomorphology, 59, 149–164, https://doi.org/10.1016/j.geomorph.2003.09.012, 2004b.

Parteli, E. J. and Herrmann, H. J.: Dune formation on the present Mars, Phys. Rev. E, 76, 041307, https://doi.org/10.1103/PhysRevE.76.041307, 2007.

Parteli, E. J., Durán, O., Bourke, M. C., Tsoar, H., Pöschel, T., and Herrmann, H.: Origins of barchan dune asymmetry: Insights from numerical simulations, Aeolian Res., 12, 121–133, https://doi.org/10.1016/j.aeolia.2013.12.002, 2014.

Pelletier, J. D.: Controls on the height and spacing of eolian ripples and transverse dunes: A numerical modeling investigation, Geomorphology, 105, 322–333, https://doi.org/10.1016/j.geomorph.2008.10.010, 2009.

Petersen, B. K., Landajuela, M., Mundhenk, T. N., Santiago, C. P., Kim, S. K., and Kim, J. T.: Deep symbolic regression: Recovering mathematical expressions from data via risk-seeking policy gradients, International Conference on Learning Representations, online, May 2021, https://openreview.net/forum?id=m5Qsh0kBQG (last access: 27 January 2023), 2019.

Piscioneri, N., Smyth, T. A., and Hesp, P. A.: Flow dynamics over a foredune scarp, Earth Surf. Proc. Land., 44, 1064–1076, https://doi.org/10.1002/esp.4555, 2019.

Psuty, N.: The coastal foredune: a morphological basis for regional coastal dune development, in: Coastal dunes: ecology and conservation, Springer, 11–27, https://doi.org/10.1007/978-3-540-74002-5_2, 2008.

Puy, A., Herzog, M., Escriche, P., Marouche, A., Oubana, Y., and Bubenzer, O.: Detection of sand encroachment patterns in desert oases. The case of Erg Chebbi (Morocco), Sci. Total Environ., 642, 241–249, https://doi.org/10.1016/j.scitotenv.2018.05.343, 2018.

Quarteroni, A., Manzoni, A., and Negri, F.: Reduced Basis Methods for Partial Differential Equations, Springer International Publishing, https://doi.org/10.1007/978-3-319-15431-2, 2016.

Rathinam, M. and Petzold, L. R.: A new look at proper orthogonal decomposition, SIAM J. Numer. Anal., 41, 1893–1925, https://doi.org/10.1137/s0036142901389049, 2003.

Richards, P. and Hoxey, R.: Appropriate boundary conditions for computational wind engineering models using the $k-\epsilon$ turbulence model, J. Wind Eng. Ind. Aerod., 46-47, 145–153, https://doi.org/10.1016/0167-6105(93)90124-7, 1993.

Richards, P. J. and Norris, S. E.: Appropriate boundary conditions for computational wind engineering: Still an issue after 25 years, J. Wind Eng. Ind. Aerod., 190, 245–255, https://doi.org/10.1016/j.jweia.2019.05.012, 2019.

Robin, N., Billy, J., Castelle, B., Hesp, P., Lerma, A. N., Laporte-Fauret, Q., Marieu, V., Rosebery, D., Bujan, S., Destribats, B., and Michalet, R.: 150 years of foredune initiation and evolution driven by human and natural processes, Geomorphology, 374, 107516, https://doi.org/10.1016/j.geomorph.2020.107516, 2021.

Roelvink, D. and Costas, S.: Coupling nearshore and aeolian processes: XBeach and duna process-based models, Environ. Modell. Softw., 115, 98–112, https://doi.org/10.1016/j.envsoft.2019.02.010, 2019.

Rudy, S., Alla, A., Brunton, S., and Kutz, N.: Data-driven identification of parametric partial differential equations, SIAM J. Appl. Dyn. Syst., 18, 643–660, https://doi.org/10.1137/18M1191944, 2019.

Rudy, S. H., Brunton, S. L., Proctor, J. L., and Kutz, J. N.: Data-driven discovery of partial differential equations, Science Advances, 3, e1602614, https://doi.org/10.1126/sciadv.1602614, 2017.

Sahoo, S. S., Lampert, C. H., and Martius, G.: Learning equations for extrapolation and control, Proceedings of the 35th International Conference on Machine Learning, PLMR, Stockholm, Sweden, July 2018, vol. 80, 4442–4450, https://proceedings.mlr.press/v80/sahoo18a.html (last access: 27 January 2023), 2018.

Salvador, M., Dedè, L., and Manzoni, A.: Non intrusive reduced order modeling of parametrized PDEs by kernel pod and neural networks, Comput. Math. Appl., 104, 1–13, https://doi.org/10.1016/j.camwa.2021.11.001, 2021.

Sauermann, G., Kroy, K., and Herrmann, H. J.: Continuum saltation model for sand dunes, Phys. Rev. E, 64, 031305, https://doi.org/10.1103/PhysRevE.64.031305, 2001.

Schatz, V. and Herrmann, H. J.: Flow separation in the lee side of transverse dunes: A numerical investigation, Geomorphology, 81, 207–216, https://doi.org/10.1016/j.geomorph.2006.04.009, 2006.

Shea, D. E., Brunton, S. L., and Kutz, J. N.: SINDy-BVP: Sparse identification of nonlinear dynamics for boundary value problems, Phys. Rev. Res., 3, 023255, https://doi.org/10.1103/physrevresearch.3.023255, 2021.

Smyth, T. A.: A review of computational fluid dynamics (CFD) airflow modelling over aeolian landforms, Aeolian Res., 22, 153–164, https://doi.org/10.1016/j.aeolia.2016.07.003, 2016.

Smyth, T. A., Jackson, D. W., and Cooper, J. A. G.: High resolution measured and modelled three-dimensional airflow over a coastal bowl blowout, Geomorphology, 177–178, 62–73, https://doi.org/10.1016/j.geomorph.2012.07.014, 2012.

Stam, J. M. T.: On the modelling of two-dimensional aeolian dunes, Sedimentology, 44, 127–141, https://doi.org/10.1111/j.1365-3091.1997.tb00428.x, 1997.

Strypsteen, G. and de Vries, S.: Simulating profile development of a new artificial dune with planted vegetation for sand nuisance mitigation, in: Coastal Sediments 2023: The Proceedings of the Coastal Sediments 2023, World Scientific, 732–738, https://doi.org/10.1142/9789811275135_0068, 2023.

Tsoar, H.: Types of aeolian sand dunes and their formation, in: Geomorphological fluid mechanics, Springer, 403–429, https://doi.org/10.1007/3-540-45670-8_17, 2001.

Tu, J. H., Rowley, C. W., Luchtenburg, D. M., Brunton, S. L., and Kutz, J. N.: On dynamic mode decomposition: Theory and applications, Journal of Computational Dynamics, 1, 391–421, https://doi.org/10.3934/jcd.2014.1.391, 2014.

van IJzendoorn, C., Hallin, C., Reniers, A., and de Vries, S.: Modeling multi-fraction coastal aeolian sediment transport with horizontal and vertical grain-size variability, J. Geophys. Res.-Earth, 128, e2023JF007155, https://doi.org/10.1029/2023JF007155, 2023.

van Westen, B.: Numerical modelling of aeolian coastal landform development, Master's thesis, Delft University of Technology, https://resolver.tudelft.nl/uuid:83344d5b-f232-4085-985c-d1ec38903254 (last access: 15 January 2024), 2018.

van Westen, B., Luijendijk, A. P., de Vries, S., Cohn, N., Leijnse, T. W., and de Schipper, M. A.: Predicting marine and aeolian contributions to the Sand Engine's evolution using coupled modelling, Coast. Eng., 188, 104444, https://doi.org/10.1016/j.coastaleng.2023.104444, 2024.

Wakes, S. J., Maegli, T., Dickinson, K. J., and Hilton, M. J.: Numerical modelling of wind flow over a complex topography, Environ. Modell. Softw., 25, 237–247, https://doi.org/10.1016/j.envsoft.2009.08.003, 2010.

Wakes, S. J., Bauer, B. O., and Mayo, M.: A preliminary assessment of machine learning algorithms for predicting CFD-simulated wind flow patterns over idealised foredunes, J. Roy. Soc. New Zeal., 51, 290–306, https://doi.org/10.1080/03036758.2020.1868541, 2021.

Walker, I. J. and Nickling, W. G.: Dynamics of secondary airflow and sediment transport over and in the lee of transverse dunes, Prog. Phys. Geog., 26, 47–75, https://doi.org/10.1191/0309133302pp325ra, 2002.

Walmsley, J. L. and Howard, A. D.: Application of a boundary-layer model to flow over an eolian dune, J. Geophys. Res.-Atmos., 90, 10631–10640, https://doi.org/10.1029/jd090id06p10631, 1985.

Walton, S., Hassan, O., and Morgan, K.: Reduced order modelling for unsteady fluid flow using proper orthogonal decomposition and radial basis functions, Appl. Math. Model., 37, 8930–8945, https://doi.org/10.1016/j.apm.2013.04.025, 2013.

Wan, Z. Y., Vlachas, P., Koumoutsakos, P., and Sapsis, T.: Data-assisted reduced-order modeling of extreme events in complex dynamical systems, PLoS ONE, 13, 1–22, https://doi.org/10.1371/journal.pone.0197704, 2018.

Weng, W. S., Hunt, J. C. R., Carruthers, D. J., Warren, A., Wiggs, G. F. S., Livingstone, I., and Castro, I.: Air flow and sand transport over sand-dunes, in: Aeolian Grain Transport, Springer Vienna, 22 pp., https://doi.org/10.1007/978-3-7091-6703-8_1, 1991.

Wippermann, F. K. and Gross, G.: The wind-induced shaping and migration of an isolated dune: A numerical experiment, Bound.-Lay. Meteorol., 36, 319–334, https://doi.org/10.1007/bf00118335, 1986.

Wu, Z., Brunton, S. L., and Revzen, S.: Challenges in dynamic mode decomposition, J. R. Soc. Interface, 18, 20210686, https://doi.org/10.1098/rsif.2021.0686, 2021.

Xiao, D., Fang, F., Pain, C., and Hu, G.: Non-intrusive reduced-order modelling of the Navier-Stokes equations based on RBF interpolation, Int. J. Numer. Meth. Fl., 79, 580–595, https://doi.org/10.1002/fld.4066, 2015.

Xiao, D., Fang, F., Pain, C. C., and Navon, I. M.: A parameterized non-intrusive reduced order model and error analysis for general time-dependent nonlinear partial differential equations and its applications, Comput. Method. Appl. M., 317, 868–889, https://doi.org/10.1016/j.cma.2016.12.033, 2017.

Yan, N. and Baas, A. C.: Parabolic dunes and their transformations under environmental and climatic changes: Towards a conceptual framework for understanding and prediction, Global Planet. Change, 124, 123–148, https://doi.org/10.1016/j.gloplacha.2014.11.010, 2015.

Yang, Y., Gu, M., Chen, S., and Jin, X.: New inflow boundary conditions for modelling the neutral equilibrium atmospheric boundary layer in computational wind engineering, J. Wind Eng. Ind. Aerod., 97, 88–95, https://doi.org/10.1016/j.jweia.2008.12.001, 2009.

Zarnetske, P. L., Hacker, S. D., Seabloom, E. W., Ruggiero, P., Killian, J. R., Maddux, T. B., and Cox, D.: Biophysical feedback mediates effects of invasive grasses on coastal dune shape, Ecology, 93, 1439–1450, https://doi.org/10.1890/11-1112.1, 2012.